PB85-245181

Ambient Ozone and Human Health
An Epidemiological Analysis
Volume 1

Resources for the Future, Inc., Washington, DC

Prepared for

Environmental Protection Agency
Research Triangle Park, NC

Aug 85

U.S. Department of Commerce
National Technical Information
NTIS

United States
Environmental Protection
Agency

Office of Air Quality
Planning and Standards
Research Triangle Park NC 27711

EPA-450/5-85-003a
August 1985

Air

PB85-245181

EPA

Ambient Ozone And Human Health: An Epidemiological Analysis

Volume I

REPRODUCED BY
NATIONAL TECHNICAL
INFORMATION SERVICE
U.S. DEPARTMENT OF COMMERCE
SPRINGFIELD, VA. 22161

TECHNICAL REPORT DATA

(Please read Instructions on the reverse before completing)

1. REPORT NO.
EPA-450/5-85-003a

2.

3. RECIPIENT'S ACCESSION NO.
PB85 245181/AS

4. TITLE AND SUBTITLE
Ambient Ozone and Human Health: An Epidemiological Analysis Volume I

5. REPORT DATE
Sept. 1985 (Date of Preparation)

6. PERFORMING ORGANIZATION CODE

7. AUTHOR(S)
Paul R. Portney and John Mullahy

8. PERFORMING ORGANIZATION REPORT NO.

9. PERFORMING ORGANIZATION NAME AND ADDRESS
Resources for the Future
1616 P Street N.W.
Washington, DC 20036

10. PROGRAM ELEMENT NO.
12A2A

11. CONTRACT/GRANT NO.
68-02-3583

12. SPONSORING AGENCY NAME AND ADDRESS
U.S. Enviornmental Protection Agency
Office of Air Quality Planning and Standards (MD-12)
Research Triangle Park, NC 27711

13. TYPE OF REPORT AND PERIOD COVERED
Final report

14. SPONSORING AGENCY CODE
OAQPS

15. SUPPLEMENTARY NOTES
Project Officer: Thomas G. Walton

16. ABSTRACT

This report is the first volume of an analysis of the relationship between ozone and human health benefits.

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group
Benefit Analysis Air Pollution, O_3 Epidemiology		

18. DISTRIBUTION STATEMENT
Release Unlimited

19. SECURITY CLASS *(This Report)*
Unclassified

20. SECURITY CLASS *(This page)*
Unclassified

21. NO. OF PAGES
324

22. PRICE

EPA Form 2220-1 (Rev. 4-77) PREVIOUS EDITION IS OBSOLETE

ATTENTION

AS NOTED IN THE NTIS ANNOUNCEMENT, PORTIONS OF THIS REPORT ARE NOT LEGIBLE. HOWEVER, IT IS THE BEST REPRODUCTION AVAILABLE FROM THE COPY SENT TO NTIS.

AMBIENT OZONE AND HUMAN HEALTH:

AN EPIDEMIOLOGICAL ANALYSIS

Paul R. Portney and John Mullahy

Resources for the Future
1755 Massachusetts Avenue, N.W.
Washington, D.C. 20036

Volume I
Draft Final Report
September 1983

Submitted to the Economic Analysis Branch, Office of Air Quality Planning and Standards, Environmental Protection Agency, Research Triangle Park, North Carolina 27711, under contract number 68-02-3583.

DISCLAIMER

This report has been reviewed by the Office of Air Quality Planning and Standards, U. S. Environmental Protection Agency, and approved for publication as received from Resources for the Future. The analysis and conclusions presented in this report are those of the authors and should not be interpreted as necessarily reflecting the official policies of the U. S. Environmental Protection Agency.

ACKNOWLEDGEMENTS

This report and the analysis on which it is based have benefitted from the efforts of a large number of individuals. Robert Fuchsberg and his colleagues in the Division of Health Interview Statistics at the National Center for Health Statistics were extraordinarily helpful in making available to us and acquainting us with the intricacies of the Health Interview Survey. Professor Raymond Palmquist of North Carolina State University gave most generously of his time to match the individuals interviewed in the 1979 Health Interview Survey to the air pollution monitors nearest their homes.

Within the Environmental Protection Agency, we received support and advice from Allen Basala, William Hunt, Bart Ostro, as well as others in the Economic Analysis Branch, the Ambient Standards Branch, the Monitoring and Data Analysis Branch, the Health Effects Research Laboratory, and the Environmental Criteria Assessment Office. Thomas Walton, our project officer, not only assisted us with countless administrative and procedural details but also made valuable substantive suggestions at each stage of our research. V. Kerry Smith, Raymond Palmquist, Duncan Holthausen, Jan Laarman, and W. Kip Viscusi--all acting as consultants to EPA--also provided valuable comments on our analysis.

A number of people knowledgeable about air pollution health effects assisted us in the early stages of our research by identifying air pollutants that might plausibly be associated with certain types of acute

and chronic illness. We are most grateful in this regard to Gilbert Omenn, Jane Koenig and Michael Morgan of the Department of Environmental Health at the University of Washington, Michael Lebowitz of the College of Medicine at the University of Arizona, and Alice Whittemore of the Stanford University School of Medicine. John Spengler of the Harvard University School of Public Health was also very helpful in this regard. We also benefitted from comments received during seminars at the University of California at Berkeley, Harvard University, Johns Hopkins University, the University of Washington and Washington State University.

Finally, we owe several of our colleagues at Resources for the Future a debt of thanks. Michael Hazilla wrote the programs which summarized the mass of air quality data with which we worked and matched it to the periods for which had individuals' health histories. Richard Carson provided expert research assistance in the early stages of the project. Among those who commented on early versions of the chapters, Robert Frank deserves special mention. He read and commented extensively on both the form and substance of every chapter. William Vaughan, Clifford Russell, Winston Harrington and Allen Kneese also provided very helpful comments on individual chapters. Margaret Parr-Recard, Pat Flynn, D.J. Curran, Anne Farr and others worked very hard typing and helping produce this draft report. We are grateful to them all for their help.

It goes without saying that none of the individuals mentioned above bears any responsibility for the contents of this report. That is the responsibility of the authors alone.

TABLE OF CONTENTS

Page

VOLUME I

CHAPTER 1. INTRODUCTION

1.1 General Background 1-1
1.2 Economic Valuation and Physical Effects 1-4
1.3 Identifying Air Pollution Health Effects 1-5
1.4 The RFF Study 1-15
1.5 Summary and Plan of this Report 1-20

CHAPTER 2. DATA

2.1 Overview 2-1
2.2 The Health Interview Survey 2-9
  2.2.1 Main Survey 2-9
  2.2.2 Smoking Supplement 2-16
  2.2.3 Residential Mobility Supplement 2-19
  2.2.4 Documentation and Cross Referencing of HIS Data 2-21
2.3 Air Pollution Data 2-22
  2.3.1 SAROAD System 2-22
  2.3.2 Data Integrity and Completeness 2-24
  2.3.3 Multi-Year Averaged Data 2-37
  2.3.4 Description of the Pollution Variables 2-42
  2.3.5 Matching Individuals to Monitors 2-46
2.4 Meteorological Data 2-47
2.5 Other Data 2-51
  2.5.1 Pollen 2-52
  2.5.2 Indoor Air Pollution 2-54
  2.5.3 Paid Sick Leave 2-56
  2.5.4 The Pollutant Standard Index 2-57
  2.5.5 availability of Medical Care 2-59
2.6 Problems of Sample Selection 2-60

CHAPTER 3. METHODOLOGY

3.1 Measures of Health Status: Choice of the Dependent Variables 3-2
3.2 Explanatory Variables 3-6
3.3 Estimation Techniques 3-52

Appendix to Chapter 3: Definitions of and Procedures Used to Create Measures of Morbidity A3-1

CHAPTER 4. RESULTS AND DISCUSSION

4.1 Acute Morbidity Due to All Causes 4-5
  4.1.1 Adults 4-5
  4.1.2 Children 4-21

4.2 Acute Morbidity Due to Respiratory Disease 4-28
    4.2.1 Adults 4-31
    4.2.2 Children 4-55
4.3 Chronic Respiratory disease 4-59
    4.3.1 Adults 4-61
    4.3.2 Children 4-81
4.4 Cardiovascular and Other Chronic Diseases 4-85
4.5 Analysis of Multicollinearity 4-89

CHAPTER 5. ALTERNATIVE OZONE CONCENTRATIONS AND CHANGES IN MORBIDITY: ILLUSTRATIVE ESTIMATES
5.1 Ozone-related Changes in Morbidity 5-2
    5.1.1 Acute Morbidity 5-3
    5.1.2 Chronic Morbidity 5-6
5.2 Acute and Chronic Illness: Feedback Effects 5-8
5.3 Relative Risks 5-11
5.4 Monetary Benefit Estimates 5-18

VOLUME II

APPENDIX A. The Health Interview and Smoking Supplement

APPENDIX B. Health Interview Survey

APPENDIX C. Valuing the Benefits of Improved Human Health

Chapter 1

INTRODUCTION

1.1 General Background

This report presents the methodology and results of a study undertaken by Resources for the Future (RFF) for the Economic Analysis Branch of the Environmental Protection Agency's Office of Air Quality Planning and Standards (OAQPS). The study was designed to identify the benefits in the form of improved human health associated with possible alternative air quality standards for ozone.

One act of Congress and one executive order provide the impetus for this study. In 1970, Congress amended the Clean Air Act of 1963 in such a way as to fundamentally alter the approach by which air quality was pursued in the United States. Among other things, the 1970 Amendments to the Clean Air Act directed the Administrator of the Environmental Protection Agency (EPA) to establish two kinds of National Ambient Air Quality Standards (NAAQS) for a handful of common (or "criteria") air pollutants. Primary standards were those designed to protect human health with an "adequate margin of safety." Secondary standards were also to be issued where appropriate; these were to be set at levels which afforded protection to vegetation, aquatic ecosystems, visibility, and other "welfare" considerations (where "welfare" refers to non-health concerns, rather than to general well-being as in economists' parlance).

Accordingly, in 1971 EPA promulgated NAAQSs for total suspended particulates, sulfur dioxide, carbon monoxide, nitrogen dioxide, and photochemical oxidants. A standard was also issued for hydrocarbons although it was intended only to assist states and localities in meeting the NAAQS for photochemical oxidants. In 1978 a NAAQS was also issued for airborne concentrations of lead. To this date, these same six pollutants are the only ones governed by NAAQSs.

Although the 1970 amendments to the Clean Air Act directed the administrator of EPA to revise the ambient standards whenever new data accumulated to so warrant, in 1977 Congress provided more specific guidance. The 1977 amendments directed EPA to review each NAAQS at least every five years and to revise the existing standard where it seemed called for, or retain the existing standard if no compelling evidence suggested that it should be changed. In 1978, in the first of these mandatory revisions, EPA proposed to change the basis of the photochemical oxidant standard to ozone (the most common photochemical oxidant), and to relax the standard from .08 parts-per-million (ppm), not to be exceeded during the highest hour each day more than once per year, to .10 ppm not to be exceeded (on an expected value basis) more than three times over a three year period. In 1979, EPA issued a final revised NAAQS for ozone. It was set at .12 ppm, to be measured in the way suggested in the proposed standard.

Because of the five-year mandatory review schedule established in the 1977 amendments to the Clean Air Act, EPA is now reviewing the scientific evidence relating ozone to human health, agricultural output, visibility impairment, and other effects. One purpose of the present study is to provide some additional information to the EPA, and specifically to the OAQPS, as part of that review.

However, the study reported here was also undertaken on account of Executive Order 12291, issued by President Reagan on February 17, 1981. That order, a successor to Executive Order 12044 issued by President Carter, directed all federal regulatory agencies in the executive branch[1] to prepare both preliminary and final Regulatory Impact Analyses (RIAs) to accompany all proposed and final "major" federal rules and regulations. These RIAs are, among other things, to describe the potential benefits expected to result from proposed or final rules, identify the potential costs, calculate the net benefits (the former minus the latter), and describe alternative approaches that might accomplish the same things as the rule in question, but at a lower cost--along with an explanation of the legal reasons why such alternatives could not be selected if, in fact, they were not.

Because the promulgation of NAAQSs have been designated as major rulemakings, Regulatory Impact Analyses have generally accompanied them. Thus, another purpose of the present study is to provide information on the human health benefits associated with possible alternative NAAQSs for ozone so as to facilitate the compilation of the overall benefits--health, agricultural, materials damage, aesthetic, and other--associated with different ozone standards. Such information, compiled from a wide variety of sources, could then be combined with data on the costs associated with these same alternatives to provide information on the overall effects expected to result from varying possible ozone standards.

1.2 Economic Valuation and Physical Effects

The original purpose of this study was to estimate, in dollar terms where possible, the value of any reduced human morbidity that might accompany reductions in ambient ozone concentrations (or the additional morbidity that might arise from a relaxation of the standard). We were to concentrate on both acute as well as chronic morbidity. That is, we were to concern ourselves with the temporary or day-to-day changes in individual health status that might result from human exposure to ambient ozone, and how these might be valued. In addition, we were to attempt to place dollar values on any long-term, or chronic, illness that might result, perhaps not so much from occasional exposures to peak ozone concentrations, but rather from prolonged exposures to ozone as well as other air pollutants.

Although the overall focus of the work reported below is still the same, the emphasis has shifted somewhat during the course of our study. Specifically, because of the availability of an extraordinary data set, it became clear during the initial phase of the RFF study that we might not only say something original about the economic valuation of ozone health effects, but might also be able to provide information about the relationship between ambient ozone concentrations and the health effects themselves. The question of economic valuation has received considerable thought during the course of our study. Appendix C to this report presents a model of household behavior from which are derived measures of the valuation of health improvements. However, more work than was originally anticipated was devoted to identifying the acute and chronic health effects that might be related to ambient exposures to ozone and other air pollutants. Thus, this report concentrates primarily on the identification and quantification of health effects resulting from exposure to ozone.

1.3 Identifying Air Pollution Health Effects

There are three primary means of ascertaining the relationship between ozone or other air pollutants and both acute and chronic morbidity. In this report we shall refer to these three approaches as toxicological, clinical, and epidemiological. Although we make use of but one of these approaches in our study, it is useful to discuss all three albeit briefly.

Under the toxicological approach, laboratory animals are exposed under carefully controlled conditions to varying levels of specific air pollutants. Dose-response relationships are estimated either by exposing otherwise identical groups of animals to different air pollution regimes and then observing differences in biological end-points (which may be functional, biochemical, structural, or behavioral), or by exposing the same animal(s) to alternative ozone or other air pollution levels at different times and observing changes in the same end-points under different regimes. The information gleaned from such animal tests is used to provide insight into the effects of these same air pollutants on human health.

As with the other approaches to health effects estimation, there are both advantages and disadvantages to the toxicological method. On the positive side, animals can be subjected to carefully controlled air pollution levels, at least some of which experimenters would be unwilling to administer to human subjects. In addition, since the lifetimes of some laboratory animals, as for example mice, may be only about two years, it is possible to observe the effects of a simulated "lifetime" exposure to air pollution in a manageable period of time. Finally, laboratory animals can be sacrificed during the course of their lives to observe directly the timing of physiological changes resulting from the exposure to air pollution that may occur anywhere within the organism.

There are, however, several serious drawbacks to toxicological studies as a means of determining human susceptibility. The most obvious of these involves the uncertainty of extrapolating the results seen in animals to human beings.[2] Researchers are unsure whether humans will respond even qualitatively to air pollution (or other environmental stimuli) in the same way as do experimental animals, much less whether the quantitative dose-response relationships estimated in animal studies are representative of likely human responses. In addition, the cost of toxicological studies sometimes necessitates that animals be given "doses" of air pollution much greater than levels to which humans might ever be exposed to make sure an effect is observed if it exists. Thus, there often arises the additional problem of extrapolating from effects observed at high doses to those observed at much lower doses, even before the inter-species extrapolations must be made.

A second approach to human health effects estimation is the clinical. At the risk of over-simplification, this involves controlled experimentation in exposure chambers or other laboratory settings using human subjects. Typically, volunteer subjects are exposed acutely--usually for periods lasting up to a few hours--to a variety of different levels of air pollution, perhaps under different exercise or other conditions as well. Their symptomatic, physiological, or in some instances biochemical responses are monitored to provide evidence about the effects of air pollution on human health.

There are several desirable features of clinical studies. First, and perhaps foremost, the dose of any pollutant(s) can be carefully controlled and measured. Thus, we can be quite sure that a certain individual was exposed to, say, .20 ppm ozone for exactly two hours and in the absence of other pollutants, because the air inflow to a chamber can be controlled quite closely. Also, because individuals have the opportunity to report symptoms as they are experienced, or because physiological changes can be measured in "real time" (i.e., contemporaneously), the likelihood of establishing a causal link is quite good. In other words, establishing such a link does not depend on an individual's recollection of a symptom sometime in the recent or distant past. Finally, if the pollutant(s) in question are thought to be particularly harmful to "sensitive individuals"--those who because of age, underlying disease, or some other distinguishing characteristic, are thought to be particularly adversely affected by the pollutant--these individuals can be selected for study.[3] In other words, the experimenter need not rely on good fortune to ensure that there are enough children, asthmatics, or other sensitive individuals in the sample group.

In spite of these attractive aspects, there are several drawbacks to the clinical approach. First, and least serious, concerns the air pollution dose that subjects receive. Although it is generally carefully measured, it is also artificial and typically is not representative of the entire mixture of air pollutants and other substances to which individuals are normally exposed in the course of their everyday lives. For example, it may not contain any of the allergens or infectious material to which we are all exposed from time to time. Thus if there are important synergistic effects between several pollutants, they may be overlooked in clinical studies.

Another problem with clinical studies concerns the numbers of participants. Perhaps on account of cost, clinical studies are generally--although not always--characterized by small numbers of subjects. It is not uncommon to find studies with fewer than ten subjects, while it is unusual to see studies in which more than fifty individuals participate. This creates two problems. First, it limits one's ability to generalize the results of a particular study to larger, potentially more diverse populations. Such generalizations will be especially difficult if the individuals studied are more sensitive in some respect than the overall population.

Small samples create another problem. For any given level of statistical significance, the smaller the number of experimental subjects the larger must be any observed effect. Since at least some air pollution health effects can be quite subtle, this means that potentially important relationships can sometimes be missed. Larger sample sizes make this less likely.

These problems pale in comparison to the most serious drawback of the clinical approach--its general inability to shed light on the possible long-term or chronic health effects associated with ozone or other air pollutants. For while subjects can be placed in artificial settings where exposures are controlled for hours or perhaps even days or weeks, legal, ethical and practical considerations prevent use of exposure regimes likely to cause persistent or irreversible damage, or that would take years to complete. Yet one important concern about air pollution is that prolonged exposures, even to levels well below those at which acute effects are thought to begin, may eventually result in chronic respiratory or other kinds of disease. Clinical studies cannot resolve this concern. In fact, air quality standards designed primarily to guard against acute health impairments associated with air pollution might still permit the onset and exacerbation of chronic illnesses that ought also to be taken into account in setting health-based air quality standards. It is here that reliance must be placed on epidemiological research.

An epidemiological study is one that looks at different groups of individuals at one point in time or the same individuals at different times (or both) and attempts to correlate differences in their health status to other differences between them (environmental, personal, and other).[4] Of course, health status is affected by much more than exposure to air pollution (or even all environmental factors) and these other factors can vary dramatically across different individuals or population groups at any one time; they can also differ for specific individuals over time--income may grow, personal habits (particularly smoking) may change, weather varies, marital status, occupation, education and a host of other factors may change and all may significantly influence health. A goal of epidemiological research is to hold constant as many of these other identifiable influences as possible. This may be done either through statistical means in the case of a heterogeneous population, or by deliberately selecting sample populations that are alike in as many ways as possible save for the characteristic(s) whose potential influences on health is to be isolated.

Epidemiological studies can be distinguished in a number of ways. One distinction revolves around the units of observation, whether individuals or aggregate populations. In the former, data are available on specific individuals--their health status, socioeconomic characteristics, environmental exposures, and so on. In aggregate epidemiology, on the other hand, a researcher tries to link variations in morbidity or mortality rates for different cities, counties, states or even countries to air pollution after attempting to control for the variation in other potentially important explanatory variables.[5] Individual data are always to be preferred where it is available because it enables one to avoid the so-called "ecological fallacy" which can arise when averaged or pooled data obscure important individual characteristics.

Another distinction that can be drawn among epidemiological studies concerns the extent to which the study is controlled or uncontrolled. In a controlled (or *ex ante*) study, the individuals or groups to be studied are identified ahead of time. They are then monitored over the course of the study period for changes in health status, personal characteristics, and so on.[6] In addition, ozone concentrations, say, are carefully monitored during the study period so that they can be linked to health status (or other dimensions of interest). In other words, such controlled or *ex ante* studies are not too unlike clinical studies save for the fact that the air pollution exposures take place in real rather than artificial settings, and for the possibly extended duration of epidemiological studies.

In uncontrolled (or *ex post*) epidemiological studies, the researcher may examine the health status of individuals or groups for whom health status, other information, and air pollution exposures are known or can be constructed, even though some or all of this information may have been collected for reasons completely unrelated to air pollution or other kinds of epidemiology.[7] In either controlled or uncontrolled studies, the quality of the results depend critically upon the breadth and reliability of the data on both the variables of interest (exposure to air pollution and health status) as well as the potentially confounding influences.

There are a number of desirable features to the epidemiological approach. Foremost among these is that, in principle, epidemiological studies have the capability of shedding light on any possible chronic health effects associated with air pollution or other factors. If one could identify individuals who have lived in one location their whole lives, if one had reliable measures of their health status, if one knew the levels of ozone and other air pollutants to which they were exposed over this period, and if one could control for the many other influences on health, then one could draw some (cautious) inferences about the link between air pollution and health. Obviously, these are very big "ifs" and the likelihood of their being satisfied is discussed below. Nevertheless, we emphasize again that, unlike clinical studies, epidemiological studies can in principle make an important contribution to the study of air pollution and chronic illness.

Another advantage to epidemiological studies is that, in general, they include larger numbers of individuals than typically participate in clincical studies. This is not inherent in the nature of such studies, but it is generally the case. Such large numbers make it more likely that any statistically significant relationships will be identified--be they between air pollution and health status or between other dependent and independent variables.

There is a final, less important advantage to epidemiological studies. By definition, such studies are based on ambient or "real world" exposures to air pollution and other environmental factors, rather than to the artificial exposures administered in clinical studies. Thus, if the exposure data are measured correctly, they accurately (indeed definitionally) reflect the conditions under which people actually receive doses of pollution. Such studies avoid the problems that may arise if laboratory settings induce by themselves certain behavioral or physiological changes.

Like the toxicological and clinical approaches, however, the epidemiological approach is not without problems. It, too, has drawbacks that can seriously limit the confidence one can place in the results of epidemiological studies. The most serious of these concerns the measurement of individual exposures to any particular air pollutant. In contrast to clinical studies, where concentrations can be carefully controlled, epidemiological studies typically rely for exposure data on readings from ambient air pollution monitors in the areas where individuals live, work and recreate. In some cases, in fact, no monitoring data may be available from the area in which one or more individuals live. And even where data are available for a particular region, they may be suspect for a number of reasons. Since individuals may in the course of their daily activities travel all about a particular metropolitan area, one must decide whether to characterize air pollution dose by the monitor nearest their homes or offices, or take some kind of weighted average based on relative amounts of time spent in various locations.

Often no data whatsoever are available about such activity patterns, so that even in the presence of comprehensive ambient monitoring data, it would be very difficult to create a weighted exposure index. Complicating the estimate of exposure still further is the matter of indoor versus outdoor exposure to air pollution.[8] It is now known, for example, that individuals may be exposed to much higher peak concentrations of nitrogen dioxide, particulate matter, and carbon monoxide in indoor than in outdoor environments even though these are three pollutants for which ambient standards exist. In addition, individuals may be exposed to radon, formaldehyde, carbon dioxide and other air pollutants in indoor environments, and these indoor concentrations may make it difficult to isolate the effects on health of ambient air quality. (Because this study deals specifically with the health effects associated with ozone, it is noteworthy that ozone is one of the air pollutants that is almost non-existent in indoor environments. With the exception of certain occupational settings (for example, welding), almost all exposures result from contact with ambient air.)

The seriousness of the problem of estimating exposure varies between epidemiological studies. In some studies the nearest air pollution monitors are so far removed from the subjects, or their data so unreliable, that little confidence can be placed in the results. In other studies, however, monitoring is done nearby and very carefully. In these cases, exposures may be well represented. Nevertheless, ambient monitoring data can be no more than a proxy for actual total exposure to ozone and other air pollutants.

Another problem, about which less will be said, concerns the controls in epidemiological studies for variables other than air pollution. Clearly, the greater the extent of controls for such variables, the greater the confidence that can be placed in any air pollution health effect that is identified. The extent of these controls varies markedly between studies--in some, such obvious and potentially confounding influences as personal smoking habits, pre-existing respiratory disease, or exposures to other air pollutants are missing. In other studies, earnest efforts are made to account for these and other much more subtle influences on health. As in the case of exposure data in epidemiological studies, the severity of the "control problem" depends on the specific measures taken in each study to address the problem.

1.4 The RFF Study

The study reported here is an epidemiological one. It combines four different types of data to shed light on acute and chronic human morbidity that may be associated with ozone, other criteria (and some non-criteria) air pollutants, and other possible contributory factors.

As described in detail below, the data on acute and chronic morbidity as well as socioeconomic characteristics are specific to individuals rather than to population groups. Thus, ours is an individual (or micro-) epidemiological study. The individual health and socioeconomic data for this study come from the 1979 Health Interview Survey (HIS) of more than 110,000 individuals, a survey that is conducted annually by the National Center for Health Statistics (NCHS) in the Department of Health and Human Services (HHS). The survey is reproduced in its entirety in Appendix A to this report.

The health data are of two types. First, information about acute illness was elicited from all individuals surveyed, information that was specific to the two-week period ending the Sunday night before the week of the interview. Individuals were asked about physical conditions that confined them to bed for most of a day, prevented normal activity (work or school) without requiring confinement to bed, or forced them to restrict their normal activity in some way even though it was not serious enough to force them to stay in bed or miss work or school. Individuals were also asked about doctor visits during the two-week period as well as short-stay hospital episodes, although these are not analyzed in our report.

In addition to information about acute health status during a specific two week period, invididuals participating in the 1979 HIS were also asked whether they had any chronic health impairments that limited their activities. Adult respondents could also identify chronic illnesses or impairments through their responses to a separate section of the 1979 HIS. Both acute and chronic impairments were identified by cause and recorded with the International Classification of Disease (ICD) codes.

One of the advantageous features of the HIS is the form of the questions about acute health status. Because such concepts as a "work loss day" or a "bed disability day" are easily understandable, individuals may be able to value the prevention of such occurrences without much difficulty. This is important in a study like this one--one aim of which is to translate physical improvements in health into dollar terms--because these individual valuations are the basis of benefit calculations in welfare economics.

This would be difficult to accomplish if each of the respondents in the HIS were given a test of lung function as part of the examination. In such a case, we would in our statistical analysis be identifying a link between ozone concentrations and lung function. If we calculated a change in the latter for a given change in ozone, the resulting effect--although important to a clinical appraisal--would not lend itself to valuation as easily as a practical endpoint like one extra or fewer day spent sick in bed. Thus, from the standpoint of applied welfare economics the HIS elicits health information in a useful form.

The socioeconomic data collected for each individual in the HIS include age, race, sex, income, education, occupation, industry of employment, marital status, history of military service, as well as other individual and household-specific characteristics discussed below.

In addition to the health and socioeconomic information that is elicited each year as part of the HIS, several supplements to the 1979 survey made it particularly useful for epidemiological purposes. Specifically, the 1979 HIS contained a supplement that went to one-third of all the adults interviewed (26,271 of a total of 79,743 adults) which provided detailed data on lifetime smoking history, including the tar and nicotine content of the brands most commonly smoked. (The smoking supplement is included in Appendix A.) This is of great importance if one is interested in examining respiratory and cardiovascular disease.

Also, the 1979 HIS included a supplement (again to one-third of all adults surveyed) designed to provide detailed information on residential histories. This information is important in examining the possible links between air pollution and chronic respiratory disease, since it makes it possible to identify individuals who have lived in their present location for a long period of time. Supplemental questions were also asked about eye care and home health care in the 1979 HIS.

To this rich set of individual health, socioeconomic, smoking and residential history data, RFF added three other types of data. The most important of these were data on air quality in the United States in 1979, consisting of all the hourly or 24-hour readings from the air pollution monitors that reported to EPA's SAROAD file (Storage and Retrieval of Aerometric Data) in 1979, as well as summary measures of this data. We also merged with the individual HIS data, other annual average pollution data dating back in some cases as far as 1974 (again from EPA's SAROAD network). Thus, we used air pollution data that were contemporaneous with the period for which we had information on individuals' acute morbidity (as opposed to using annual average air pollution to proxy exposure during a specific part of the year). In addition, we had annual average air pollution data from the urban areas in which the individuals in our sample resided (for both 1979 as well as previous years in many cases) to assist in exploring the possible role of air pollution in chronic disease etiology.

In addition to this contemporaneous matching of individuals to air quality data, we also tried for a finer spatial resolution of air quality data. Using a procedure discussed in Chapter 2, we matched those individuals in the 1979 Health Interview Survey who lived in Standard Metropolitan Statistical Areas (SMSAs) to the nearest ten monitors for each of eight pollutants--ozone, total suspended particulate matter, nitrogen dioxide, total oxides of nitrogen, sulfur dioxide, carbon monoxide, lead and sulfates. In addition, we knew the distance between the centroid of the census tract in which an individual resided and the location of these monitors. Thus, although we were free to use SMSA-wide averages to represent air pollution exposures at any point in time, or construct our own distance-weighted average exposures, we had the capability to match each individual in our sample to the air pollution monitor(s) nearest his or her home and use that as a measure of exposure. The latter was the approach used in our study.

We also added to our data base information on weather conditions specific to the two-week period for which we had acute health data for a particular individual, as well as averaged over the year 1979. These data came from the National Oceanographic and Atmospheric Administration (NOAA) and were measured at one or more weather stations in each SMSA, generally at airports. Included in the weather data were observations on temperature (both central tendencies and extremes), precipitation, humidity and other measures.

Finally, we compiled a third data set, one much more heterogeneous than the air pollution or meteorlogical data. This included a number of potential influences on acute and chronic health status which were unrelated to socioeconomic status, air pollution concentrations, or weather conditions. Included in this category were: ragweed concentrations for a subset of the metropolitan areas from which the individuals in our sample were drawn; the probability that an individual came from a household where piped or bottled natural gas was the predominant cooking fuel (because of its potential for indoor air pollution); a measure of the availability of doctors in the area in which an individual lived; an estimate of the annual amount of paid sick leave to which an individual was entitled (because sick leave may affect work loss); as well as a number of other unrelated but potentially important factors.

A final observation about the RFF study concerns our reliance on medical and epidemiological experts. Every study should try to build upon and learn from the advances made in previous studies. The present study is no exception. We have tried to draw on the best features of previous epidemiological, clinical and toxicological studies. These were useful in helping identify possible explanatory variables and in constructing hypotheses about the kinds of acute and chronic health impairments expected to result from short-term peak exposures or longer-term exposures to ozone and other air pollutants. Of particular assistance were a number of doctors and/or epidemiologists who identified specific diseases and symptoms as being likely to result from exposures to certain air pollutants. Although they were not acting as paid consultants to our study, their assistance was most important in the construction of the hypotheses we tested below.

1.5 Summary and Plan of this Report

The study described in this report was an ambitious one. It used an unusually large and detailed individual data set; combined these individual data with air pollution and meteorological data that were as finely resolved in time and space as the SAROAD network, the NOAA system and the HIS permitted; included explanatory variables which, to our knowledge, were not used before; experimented with a number of functional forms, model specifications, and definitions or measurements of individual health status; and reported the results of a great many sensitivity tests designed to shed light on the confidence one can place in the results reported below.

Nevertheless, it is important to keep this study in its proper persepctive. Because it is an epidemiological study, it suffers from some of the problems identified earlier. For instance, although individuals were matched to the air pollution monitors nearest their homes, we still lacked information on their daily activity patterns. Thus, we could not characterize precisely their exposure to ambient pollutant concentrations. Even if we had, potential exposure to indoor pollutants would still have complicated the interpretation of our results. Until personal monitors are a practical solution, however, some imperfection in characterizing exposure and dosage must be expected. Thus, our results must be interpreted carefully in view of the imperfect exposure measures we used.

Second, in spite of the detailed socioeconomic data contained in the HIS, we lacked data on certain individual characteristics or habits which may play an important role in determining acute or chronic health status. For instance, individuals surveyed by NCHS were asked nothing about their exercise or sleeping habits, dietary practices, or alcohol consumption, even though all of these may influence health in one way or another. Similarly, while we had information on individuals' occupations and the industries in which they are employed, it would have been useful to have data on exposures to hazardous substances in the workplace. In some of these instances, we constructed crude proxies for these factors. But these proxies were second-best alternatives to more detailed information. Throughout the study we point to respects in which the estimates could be improved by additional information. Nevertheless, we must emphasize at the outset that we lacked perfect control measures.

Because of these (almost inevitable) limitations, as well as others identified below, the results of this study must be carefully interpreted. It is important when reading the following chapters to pay as much attention to the caveats about our estimates as to the estimates themselves. Similarly, one must pay close attention to the sensitivity of a particular result to the form of the estimating equation, the specification of the model, the measurement of the dependent and independent variables, and the sample over which the results were estimated.

No single epidemiological or other kind of study can ever come up with the "right" answer. Thus, the results of this study must be viewed alongside those from other epidemiological, clinical, and toxicological studies. One can have more confidence in results emerging from a variety of different kinds of studies than in results specific to a particular study or class of studies.

The plan of this report is as follows. Chapter 2 discusses the data used in this report, and is divided into five sections. The first of these deals with the Health Interview survey, and goes into some detail about the main survey (in particular the definitions of acute and chronic morbidity), the smoking supplement, and the supplement on residential mobility. A second section deals with the air pollution data used in the analysis. Included is a discussion of the EPA air pollution monitoring data, the completeness and quality assurance methods to used to screen these data, the procedure we used to create a measure of pollution exposure over time, our approach to naming the pollution variables in the regressions below, and the procedure used to match individuals in the HIS to the air pollution monitoring sites. Two sections then follow on the meteorological and other data employed in our analysis and the purposes for which they are used. A final section takes up the question of sample selection bias in the method used to create our final sample of adults and children.

Chapter 3 of the report presents the methodology used in our analysis of acute and chronic morbidity. Included there is: a discussion of the dependent variables used in the analysis and the way in which they were created from the HIS data tapes; a presentation of the independent variables used to help explain acute and chronic morbidity (and a discussion about possible collinearity between them); and a review of the estimating techniques employed in the analysis.

In Chapter 4 are presented the results of our study. Separate sections are devoted to the analysis of acute morbidity due to all causes, acute morbidity due specifically to respiratory disease, chronic respiratory disease, and cardiovascular and other diseases. We also discuss in Chapter 4 the diagnostic tests we have conducted to identify any collinearity that may exist between independent variables.

Finally, in Chapter 5 we present some estimates of the changes in both acute and chronic health status that might result from changes in ambient ozone concentrations in the urban areas of the U.S. These predicted changes might form the basis for estimates--in dollar terms, perhaps--of the benefits (or health costs) that might result from alternative National Ambient Air Quality Standards for ozone.

Footnotes

[1] Technically, the executive order applies to the regulatory agencies governed by 44 United States Code 3502(1), excluding those agencies specified in 44 United States Code 3502(10).

[2] See Schneiderman, Mantel, and Brown (1975).

[3] On sensitive populations, see Friedman (1981).

[4] For a general introduction to epidemiology, see Lilienfeld and Lilienfeld (1980).

[5] For example, see Lave and Seskin (1979).

[6] See Whittemore and Korn (1980), for an example.

[7] See Seskin (1979), for instance.

[8] For an introduction to and discussion of indoor air pollution and its possible effects on health, see National Academy of Sciences (1981) and the General Accounting Office (1980).

References

Friedman, Robert, Sensitive Populations and Environmental Standards (Washington, D.C.: The Conservation Foundation), 1981.

General Accounting Office, "Indoor Air Pollution: An Emerging Health Problem," report no. CED-80-111 (1980).

Lave, Lester and Eugene Seskin, Air Pollution and Human Health (Baltimore, Md.: The Johns Hopkins University Press), 1977.

Lilienfeld, Abraham and David Lilienfeld, Foundations of Epidemiology (New York: Oxford University Press), 1980.

National Academy of Sciences, Indoor Pollutants (Washington, D.C.: National Academy Press) 1981.

Schneiderman, Marvin, Nathan Mantel, and Charles Brown, "From Mouse to Man--or How to Get from the Laboratory to Park Avenue and 59th Street," Annals of the New York Academy of Sciences, vol. 246 (1975) pp. 237-248.

Seskin, Eugene, "An Analysis of Some Short-Term Health Effects of Air Pollution in the Washington, D.C. Metropolitan Area," Journal of Urban Economics, vol. 6 (1979) pp. 275-91.

Whittemore, Alice and Edward Korn, "Asthma and Air Pollution in the Los Angeles Area," American Journal of Public Health, vol. 70 (1980), pp. 687-96.

Chapter 2

DATA

2.1 Overview

The purpose of the data used in this study is to assist in determining the relationships, if any, between individuals' acute and chronic morbidity and their exposure to air pollutants, while controlling for other factors. As Chapter 1 pointed out, this is not an easy task given imperfections in the data, problems in measuring exposure, and an array of other potentially confounding factors. Nonetheless, we believe the data used in this study improve upon those used in prior epidemiologic investigations in several important respects.

The basic approach in this study is the following. A survey of individuals' health and socioeconomic status serves as the foundation of the data base. With the help of the National Center for Health Statistics we matched each of those interviewed in the 1979 Health Interview Survey (HIS) to the nearest ten monitors for each of eight different pollutants by assuming that each household lived at the centroid of the census tract in which its house or apartment was located. This was accomplished via a computer algorithm that uses input data on the geographic coordinates of census tracts from the Urban Atlas Files of the Bureau of Census, as well as data on the geographic coordinates of air pollution monitors from the EPA's Directory of Air Quality Monitoring Sites.[1] Details of the procedure are described later in this chapter.

Because of the need to preserve confidentiality, we could not be given either the address or census tracts of individuals in the HIS. To get around this problem we provided the NCHS with a tape which matched each census tract in every SMSA in the United States with the nearest ten monitors for each of eight different pollutants. Using the census tract identifiers on their "master" HIS data tape, NCHS matched individuals to monitors and included this monitoring information on the HIS tape they sent us (a tape which did not include the census tract identifiers). Matching weather stations and other area-specific data was accomplished at the same time.

Finally, pollution data gathered from these stations were summarized and combined with the individual health and socioeconomic data. The pollution data were averaged over three periods: the two-week period in 1979 for which data were available on each individual's acute morbidity; the entire year 1979; and, in some cases, the period 1974-79 (the uses of these data are explained below). A similar procedure was used for matching summarized meteorological data to the individuals interviewed by NCHS.

Two distinct groups of individuals were analyzed in this study--men and women aged seventeen and above (designated "adults"), and children and adolescents under seventeen (designated "children"). Both groups were drawn from the 1979 Health Interview Survey. This was a survey of 110,530 individuals of all ages and from all geographical regions of the United States. In forming the "adults" data set, the original sample of 110,530 was first cut by requiring that all adults had been given the supplemental smoking survey, since estimation of the effects of air pollution on respiratory and cardiovascular health was the primary focus of the study. The smoking supplement was administered to approximately one-third of the adults in the main survey, or 26,271 individuals. This number was reduced further because air pollution data were lacking for a portion of these adults (described below). This reduced the number of adults to 14,446. Finally, five individuals with internally inconsistent data were deleted, bringing the number of adults to be studied to 14,441. Table 2-1 indicates the distribution across U.S. metropolitan areas of those in the final RFF sample of adults.

The data set for "children" was created the same way, although no observations were lost for lack of smoking data since the smoking supplement was only administered to those aged seventeen or above. The number of children used in the analysis below is 15,711. Later in this chapter we discuss the possible bias introduced by our sample selection procedures for adults and children.

Table 2-1. Residence of Individuals in RFF's Adults Data Set

SMSA Number	Location of SMSA	Number of Individuals	Percent of Total
0080	Akron, Ohio	65	0.45
0160	Albany-Schenectady-Troy, NY	82	0.57
0200	Albuquerque, NM	31	0.22
0240	Allentown-Bethlehem-Easton, PA	69	0.48
0320	Amarillo, TX	38	0.26
0360	Anaheim-Santa Ana-Garden Grove, CA	170	1.18
0400	Anderson, IN	65	0.45
0460	Appleton-Oshkosh, WI	30	0.21
0480	Ashville, NC	64	0.44
0520	Atlanta, GA	146	1.01
0600	Augusta, GA-SC	26	0.18
0640	Austin, TX	37	0.26
0680	Bakersfield, CA	40	0.28
0720	Baltimore, MD	199	1.38
0760	Baton Rouge, LA	34	0.24
0800	Bay City, MI	34	0.24
0840	Beaumont-Port Arthur-Orange, TX	30	0.21
0920	Biloxi-Gulfport, MS	50	0.35
0960	Binghamton, NY-PA	38	0.26
1000	Birmingham, AL	93	0.64
1120	Boston, MA	276	1.91
1160	Bridgeport, CT	37	0.26
1170	Bristol, CT	5	0.04
1200	Brockton, MA	25	0.17
1240	Brownsville-Harlingen-San Benito, TX	36	0.25
1280	Buffalo, NY	145	1.00
1320	Canton, OH	42	0.29
1440	Charleston, SC	21	0.15
1520	Charlotte, NC	47	0.33
1600	Chicago, IL	703	4.87
1640	Cincinnati, OH-IN-KY	157	1.09
1680	Cleveland, OH	218	1.51
1720	Colorado Springs, CO	49	0.34
1740	Columbia, MO	45	0.31
1760	Columbia, SC	32	0.22
1800	Columbus, GA-AL	19	0.13
1840	Columbus, OH	92	0.64
1880	Corpus Christi, TX	42	0.29
1920	Dallas-Ft. Worth, TX	193	1.34
1960	Davenport-Rock Island-Moline, IA-IL	40	0.28
2000	Dayton, OH	89	0.62
2080	Denver, CO	124	0.86

SMSA Number	Location of SMSA	Number of Individuals	Percent of Total
2120	Des Moines, IA	26	0.18
2160	Detroit, MI	407	2.82
2240	Duluth-Superior, MN-WI	16	0.11
2320	El Paso, TX	49	0.34
2360	Erie, PA	31	0.22
2400	Eugene, OR	29	0.20
2440	Evansville, IN-KY	22	0.15
2480	Fall River, MA-RI	22	0.15
2520	Fargo-Moorhead, MN-ND	48	0.33
2600	Fitchburg-Leominster, MA	14	0.10
2640	Flint, MI	57	0.40
2680	Ft. Lauderdale-Hollywood, FL	73	0.51
2760	Ft. Wayne, IN	20	0.14
2840	Fresno, CA	51	0.35
2880	Gadsden, AZ	42	0.29
2920	Galveston-Texas City, TX	40	0.28
2960	Gary-Hammond-East Chicago, IN	51	0.35
3000	Grand Rapids, MI	53	0.37
3040	Great Falls, MT	34	0.24
3080	Green Bay, WI	52	0.36
3120	Greensboro-Winston Salem, NC	66	0.46
3160	Greensville, SC	45	0.31
3240	Harrisburg, PA	34	0.24
3280	Hartford, CT	70	0.49
3320	Honolulu, HI	69	0.48
3360	Houston, TX	230	1.59
3400	Huntington-Ashland, OH-WVA-KY	32	0.22
3440	Huntsville, AL	33	0.23
3480	Indianapolis, IN	103	0.71
3520	Jackson, MI	37	0.26
3560	Jackson, MS	24	0.17
3600	Jacksonville, FL	58	0.40
3640	Jersey City, NJ	61	0.42
3680	Johnstown, PA	24	0.17
3720	Kalamazoo, MI	39	0.27
3760	Kansas City, KS-MO	132	0.91
3840	Knoxville, TN	39	0.27
3880	Lafayette, LA	45	0.31
4000	Lancaster, PA	44	0.31
4040	Lansing-East Lansing, MI	46	0.32
4120	Las Vegas, NV	35	0.24
4160	Lawrence-Haverhill, MA-HN	30	0.21
4280	Lexington, KY	41	0.28
4320	Lima, OH	35	0.24

SMSA Number	Location of SMSA	Number of Individuals	Percent of Total
4400	Little Rock, N. Little Rock, AR	38	0.26
4440	Lorain-Elyria, OH	27	0.19
4480	Los Angeles-Long Beach, CA	754	5.22
4520	Louisville, IN-KY	88	0.61
4560	Lowell, MA	19	0.31
4640	Lynchburg, VA	38	0.26
4720	Madison, WI	36	0.25
4760	Manchester, NH	13	0.09
4800	Mansfield, OH	46	0.32
4880	McAllen-Pharr-Edinburg, TX	37	0.26
4920	Memphis, TN-AR	76	0.53
4960	Meridian, CT	4	0.03
5000	Miami, FL	142	0.98
5040	Midland, TX	41	0.28
5080	Milwaukee, WI	151	1.05
5120	Minneapolis-St. Paul, MN	180	1.25
5160	Mobile, AL	38	0.26
5360	Nashville, TN	68	0.47
5400	New Bedford, MA	17	0.12
5440	New Britain, CT	11	0.08
5480	New Haven-West Haven, CT	39	0.27
5560	New Orleans, LA	116	0.80
5601-08	New York, NY	1,184	8.20
5640	Newark, NJ	210	1.45
5680	Newport News-Hampton, VA	30	0.21
5720	Norfolk-Portsmouth, VA-NC	64	0.44
5760	Norwalk, CT	9	0.06
5880	Oklahoma City, OK	72	0.50
5920	Omaha, NE-IA	43	0.30
5960	Orlando, FL	41	0.28
6000	Oxnard-Ventura, CA	38	0.26
6040	Patterson-Clifton-Passaic, NJ	178	1.23
6080	Pensacola, FL	31	0.22
6120	Peoria, IL	33	0.23
6160	Philadelphia, PA-NJ	483	3.35
6200	Phoenix, AZ	88	0.61
6280	Pittsburgh, PA	279	1.93
6320	Pittsfield, MA	11	0.08
6400	Portland, ME	47	0.33
6440	Portland, OR-WA	70	0.49
6480	Providence-Pawtucket-Warwick, RI	98	0.68
6600	Racine, WI	39	0.27
6640	Raleigh-Durham, NC	16	0.11
6680	Reading, PA	32	0.22

SMSA Number	Location of SMSA	Number of Individuals	Percent of Total
6760	Richmond, VA	27	0.19
6800	Roanoke, VA	26	0.18
6840	Rochester, NY	104	0.72
6880	Rockford, IL	30	0.21
6920	Sacramento, CA	82	0.57
7040	St. Louis, MO-IL	217	1.50
7120	Salinas-Monterrey, CA	27	0.19
7160	Salt Lake City, UT	52	0.36
7240	San Antonio, TX	96	0.67
7280	San Bernardino-Riverside-Ontario, CA	134	0.93
7320	San Diego, CA	146	1.0
7360	San Francisco-Oakland, CA	328	2.27
7400	San Jose, CA	61	0.42
7480	Santa Barbara, CA	43	0.30
7500	Santa Rosa, CA	48	0.33
7520	Savannah, GA	44	0.31
7600	Seattle-Everett, WA	170	1.18
7680	Shreveport, LA	29	0.20
7800	South Bend, IN	28	0.19
7840	Spokane, WA	29	0.20
8000	Springfield-Chicopee-Holyoke, MA-CT	57	0.40
8040	Stamford, CT	32	0.22
8080	Steubensville-Weirton, OH-WVA	14	0.10
8120	Stockton, CA	35	0.24
8160	Syracuse, NY	70	0.49
8200	Tacoma, WA	37	0.26
8280	Tampa-St. Petersburg, FL	90	0.62
8320	Terre Haute, IN	50	0.35
8360	Texarkana, TX-AR	38	0.26
8400	Toledo, OH-MI	67	0.46
8440	Washington, DC-VA-MD	36	0.25
8480	Trenton, NJ	28	0.19
8520	Tucson, AZ	39	0.27
8560	Tulsa, OK	51	0.35
8680	Utica-Rome, NY	31	0.22
8720	Vallejo-Napa, CA	21	0.15
8840	Washington, DC-MD-VA	258	1.79
8880	Waterbury, CT	26	0.18
8920	Waterloo, IA	37	0.26
8960	West Palm Beach, FL	37	0.26
9000	Wheeling, OH-WVA	13	0.09
9040	Witchita, KS	35	0.24
9120	Wilkes Barre-Hazelton, PA	38	0.26
9160	Wilmington, DE-NJ-MD	51	0.35

SMSA Number	Location of SMSA	Number of Individuals	Percent of Total
9200	Wilmington, NC	41	0.28
9240	Worcester, MA	29	0.20
9280	York, PA	39	0.17
9320	Youngstown-Warren, OH	58	0.40

2.2 The Health Interview Survey

2.2.1 Main Survey

The 1979 Health Interview Survey was administered by the National Center for Health Statistics (NCHS) of the Public Health Service of the U.S. Department of Health and Human Services. The following general description may be useful:

> The National Health Interview Survey utilizes a questionnaire that obtains information on personal and demographic characteristics, illnesses, injuries, impairments, chronic conditions, and other health topics. The population covered by the sample for the National Health Interview Survey is the civilian, noninstitutionalized population of the United States living at the time of the interview. The sample does not include members of the Armed Forces or U.S. nationals living in foreign countries.
>
> The sampling plan of the survey follows a multistage probability design which permits a continuous sampling of the civilian noninstitutionalized population of the United States. The sample is designed in such a way that the sample of households interviewed each week is representative of the target population and that weekly samples are additive over time. This feature of the design permits both continuous measurement of characteristics of samples and more detailed analysis of less common characteristics and smaller categories of health-related items. The continuous collection has administrative and operational advantages as well as technical assets since it permits fieldwork to be handled with an experienced, stable staff.
>
> The overall sample was designed so that tabulations can be provided for each of the four major geographic regions and for selected places of residence in the United States.
>
> The first stage of the sample design consists of drawing a sample of 376 primary sampling units (PSU's) from approximately 1,900 geographically

> defined PSU's. A PSU consists of a county, a small group of contiguous counties, or a standard metropolitan statistical area. The PSU's collectively cover the fifty states and the District of Columbia.
>
> With no loss in general understanding, the remaining stages can be combined and treated in this discussion as the ultimate stage. Within PSU's, then, ultimate stage units called segments are defined in such a manner that each segment contains an expected six households. Three general types of segments are used.
>
> - Area segments which are defined geographically
> - List segments, using 1970 census registers as the frame
> - Permit segments, using updated lists of building permits issued in sample PSU's since 1970
>
> Census address listings were used for all areas of the country where addresses were well defined and could be used to locate housing units. In general the list frame included the larger urban areas of the United States from which about two-thirds of the NHIS sample was selected.
>
> The usual NHIS sample consists of approximately 12,000 segments containing about 50,000 assigned households, of which 9,000 were vacant, demolished, or occupied by persons not in the scope of the survey. The 42,000 eligible occupied households yield a probability sample of about 111,000 persons.[2]

In 1979, there was a total noninterview rate of approximately 3.9 percent, consisting of individuals refusing to respond or who could not be located at home after repeated calls.

This is a cursory description of the HIS; the interested reader is referred to more detailed information available elsewhere.[3]

The HIS provides a rich description of the socioeconomic characteristics of those surveyed. Each year information is sought about individuals' health status and household and family characteristics, educational attainment, income, occupation, number of hours worked, and many other subjects of potential interest to researchers. Furthermore, the 1979 HIS contained detailed supplemental surveys on smoking behavior and residential history, each going to one-third of the adults surveyed (the selection of subsamples produced considerable overlap between those who received the smoking survey and those who received the residential migration survey.) Appendix A to this report presents the 1979 Health Interview Survey in its entirety, as well as the smoking and residential history supplements.

It is for the provision of detailed information about individuals' health status, however, that the HIS was selected as the foundation of our data base. It is designed to provide a detailed picture of interviewees' health status at the time of the interview, rather than broader and more general information. Because the data are self reported, response biases are possible with respect to health and other types of information that could confound estimation results.[4] Further, because the HIS was not designed solely for the purposes of epidemiologic research, but rather to provide a detailed picture of health in the U.S., it does not elicit certain information one would collect in a survey designed explicitly for epidemiological purposes.

Because of the size of the HIS sample, however, and the details available on individuals' health and other characteristics, the HIS is a very rich data base.

As discussed above, it is difficult to estimate the determinants of acute and chronic morbidity. In fact it is hard to define sickness or health, much less summarize it by a quantitative or qualitative index that can be utilized in statistical estimation.[5] The problem is compounded when using the HIS because health measures are self-reported rather than physician-diagnosed. We now turn to a discussion of the health status measurements available in the HIS and utilized in the present study.

It is not uncommon in health surveys to ask individuals if they consider their health to be excellent, good, fair, or poor. The "EGFP" question, as it has come to be known, is included in the HIS. The response to such a question conveys too little objective information about health, however, and two individuals in equivalent states of health might answer differently for totally unrelated reasons. Indeed, the same individual might answer the question differently--without any change in objective health status--at two different times during the same day.[6]

The HIS, however, goes much further in eliciting health information. Survey questions are aimed at ascertaining not only whether the respondent did or did not have any of a detailed set of chronic or acute conditions or impairments defined below, but also the severity of such conditions. Severity is quantified in a variety of ways, including time since onset of condition, duration of activity limitation, and degree of physical disability due to limitations. The following describes the measures of health status compiled in that survey.

Terms Relating to Conditions

Condition--A morbidity condition, or simply a condition, is any entry on the questionnaire that describes a departure from a state of physical or mental well-being. It results from a positive response to one of a series of "medical-disability impact" or "illness-recall" questions. In the coding and tabulating process, conditions are selected or classified according to a number of different criteria (such as whether they were medically attended, whether they resulted in disability, or whether they were acute or chronic) or according to the type of disease, injury, impairment, or symptom reported.

Conditions except impairments are classified by type according to the ninth revision of the International Classification of Diseases, with certain modifications adopted to make the code more suitable for a household interview survey.

Chronic Condition--A condition is considered chronic if (1) the condition is described by the respondent as having been first noticed more than three months before the week of the interview, or (2) it is one of the following conditions always classified as chronic regardless of the onset.

Tuberculosis
Neoplasms (benign and malignant)
Diseases of the thyroid gland
Diabetes
Gout
Psychoses and certain other diseases of the central nervous system
Multiple sclerosis and certain other diseases of the central nervous system
Certain diseases of the eye
Certain diseases of the circulatory system (includes rheumatic fever, hypertension, stroke, and all heart conditions)
Emphysema, asthma, hay fever, and bronchiectasis

Ulcers and certain other diseases of the esophagus, stomach, and duodenum
Hernia of abdominal cavity (includes rupture)
Gastroenteritis and colitis (with exceptions)
Calculus of kidney, ureter, and other parts of the urinary system
Diseases of the prostate
Chronic cystic diseases of the breast
Eczema and certain other dermatitis
Arthritis and rheumatism
Cyst of the bone (except jaw)
All congenital anomalies

Impairments--Impairments are chronic or permanent defects, usually static in nature, that result from disease, injury, or congenital malformation. They represent decrease or loss of ability to perform various functions, particularly those of the musculoskeletal system and the sense organs.

Onset of condition--A condition is considered to have had its onset when it was first noticed. This could be the time the person first felt sick or became injured, or it could be the time when the person or family was first told by a physician that the person had a condition of which he or she had been previously unaware.

Terms Relating to Disability

Disability--Disability is the general term used to describe any temporary or long-term reduction of a person's activity as a result of an acute or chronic condition.

Disability day--Short-term disability days are classified according to whether they are days of restricted activity, bed days, hospital days, work-loss days, or school-loss days. All hospital days are, by definition, days of bed disability; all days of bed disability are, by definition, days of restricted activity. The converse form of these statements is, of course, not true. Days lost from work and days lost from school are special terms that apply to the working and school-age populations only but these too are days of restricted activity. Hence, "days of restricted activity" is the most inclusive term used to describe disability days.

Restricted-activity day--A day of restricted activity is one on which a person cuts down on his or her usual activities for the whole of that day because of an illness or an injury. The term "usual activities" for any day means the things that a person would ordinarily do on that day. For children under school age, usual activities depend on whatever the usual pattern is for the child's day, which will in turn be affected by the age of the child, weather conditions, and so forth. For retired or elderly persons, usual activities might consist of almost no activity, but cutting down on even a small amount for as much as a day would constitute restricted activity. On Sundays or holidays, usual activities are the things the person usually does on such days--going to church, playing golf, visiting friends or relatives, or staying home and listening to the radio, reading, looking at television, and so forth. Persons who have permanently reduced their usual activities because of a chronic condition might not report any restricted-activity days during a two-week period. Therefore, absence of restricted-activity days does not imply normal health.

Restricted activity does not imply complete inactivity, but it does imply only the minimum of usual activities. A special nap for an hour after lunch does not constitute cutting down on usual activities, nor does the elimination of a heavy chore such as cleaning ashes out of the furnace or hanging the wash. If a farmer or housewife carries on only the minimum of the day's chores, however, this is a day of restricted activity.

A day spent in bed or a day home from work or school because of illness or injury is, of course, a restricted-activity day.

Bed-disability day--A day of (bed)-disability is one on which a person stays in bed for all or most of the day because of a specific illness or injury. All or most of the day is defined as more than half of the daylight hours. All hospital days for inpatients are considered to be days of bed disability even if the patient was not actually in bed at the hospital.

Work-loss day--A day lost from work is a day on which a person did not work at his job or business for at least half of his normal workday because of a specific illness or injury. The number of days lost from work is determined only for persons seventeen years of age and over who reported that at any time during the two-week period covered by the interview they either worked at or had a job or business.

School-loss day--A day lost from school is a normal school day on which a child did not attend school because of a specific illness or injury. The number of days lost from school is determined only for children six through sixteen years of age.[7]

2.2.2 Smoking Supplement

In its administration of the annual Health Interview Survey, the NCHS includes what are termed rotating supplemental surveys which are administered at less than an annual frequency. One such survey of relevance to the present study is the smoking supplemental survey included as a part of the 1979 HIS. Due to the nature of our study--where analysis of the effects of pollutants on individuals' health must necessarily be concentrated largely on respiratory conditions--the availability of detailed information on individuals' smoking histories is a necessity if the effects of air pollution on health are to be properly assessed.

Detailed data are available from the smoking supplement. Individuals' responses to questions about the following were recorded in the 1979 survey:

1. Smoking status (Never, Former, Occasional, Current)
2. Current smoking status (Occasional Smokers in 1)
3. Number of cigarettes now smoked a day
4. Age started smoking regularly
5. Number of cigarettes smoked a day at peak period
6. Last smoked regularly
7. Interval since last smoked regularly (Former smokers in 1)
8. Number of brands smoked
9. Type of filter
10. Kind of cigarette smoked (plain, menthol)
11. Package type

12. Cigarette size
13. Ever seriously attempted to quit
14. Total number of serious attempts to quit
15. Number of serious attempts in past twelve months
16. Starting time of last attempt
17. Length of time stayed off cigarettes
18. Tar level of brand smoked most
19. Nicotine level of brand smoked most

In 1979, 26,271 adults (aged seventeen and above) were surveyed about their smoking habits, this representing approximately one-third of all adults interviewed in the main survey. A facsimile of the 1979 smoking survey is included with the facsimile of the main survey in Appendix A to this report.

One of the drawbacks of the smoking supplement is that typically only one adult in each household (or two adults if there were four or more adults in residence) was surveyed about his or her smoking behavior. This makes it difficult to estimate the effects of sidestream smoking on the health of children and other resident adults. In assembling the data used in the present study, with the exception of the data on children, we have deleted those individuals for whom no smoking data were collected. In the estimation phase of the study, this presents some problems. For example, if a respondent reports him- or herself to be a nonsmoker, there exists the possibility that the individual is exposed to sidestream smoke in the house, apart from any sidestream smoke experienced while out of the house. The Surgeon General of the United States has reported that:

> Children of smoking parents had more bronchitis and pneumonia during the first year of life; and acute respiratory disease accounted for a higher number of restricted activity days (1.1 days) and bed disability days (0.8 day) in children whose families smoked than in those whose families did not. A reduction in exercise tolerance with exposure to sidestream cigarette smoke has been demonstrated in patients with angina pectoris, and a decrease in small airway function of the lung equivalent to that observed in light smokers (one to ten cigarettes a day) has been reported in adults who never smoked themselves nor lived with smokers, but who were exposed to cigarette smoking in the workplace.[8]

The results of the smoking data should be interpreted carefully since there is evidence of an increased underreporting of cigarette consumption over time. One researcher has concluded:

> Data from four major surveys, spanning the years since the Surgeon General's Report, suggest a significant reduction in rates of cigarette smoking. These data, however, conflict with production and sales data which record only a slight reduction. Explanations for this discrepancy range from problems in the surveys' methodology to increased underreporting of cigarette consumption because of both a growing awareness of the threat to health and the social undesirability of smoking.[9]

Smoking has become more stigmatized, which may account for whatever response bias exists. Future research should control for the possible errors-in-variables problem with the quantitative or qualitative smoking measures used in analysis.

Table 2-2 summarizes the smoking habits of all individuals receiving the 1979 HIS smoking supplement as well as the characteristics of the adults used in the present study. The latter are members of the former for whom air pollution data were available. It is clear that the smoking habits of our subsample are virtually identical to those of the universe of individuals answering the smoking supplement.

2.2.3 Residential Mobility Supplement

As part of the 1979 Health Interview Survey, a subset of adults was surveyed about their residential histories. They were asked about how long they had been at their present address, how many times they had moved in the past three years, how far they had moved, and when they had moved. Data were summarized for 25,519 individuals, most of whom were also given the smoking supplemental survey.

The reason for our interest in information about individuals' residential histories centers on the analysis of the possible relationship between air pollution and chronic morbidity. Previous investigations of this relationship have proceeded by examining associations between chronic disease and current air pollution concentrations in the area where respondents currently live. Current air pollution levels are taken as representative of lifetime exposures.

Table 2.2. Comparison of Smoking Habits of the General Population (as Sampled in the HIS) and Those in the RFF Subsample

	Main Survey (n = 26,271)	RFF Subsample (n = 14,441)
Smoking Status		
% Never	43.1	41.9
% Occasional	2.2	2.1
% Former	18.0	18.3
% Current	30.5	31.6
% Unknown	6.2	6.2
Number of Cigarettes Per Day, Present Smokers, Median	20	20
Age Started Smoking Regularly, Median	17	17
Tar Level of Brand Smoked Most (mgs), Median	16.9	16.9
Nicotine Level of Brand Smoked Most (mgs), Median	1.05	1.05

But on average individuals move every five years and thus live in a number of places in the course of their lifetimes. Thus, a chronic illness that may have been initiated and/or promoted by poor air quality in one area might be falsely associated with air quality in the area where the individual is living at the time of the study. Data on residential histories can be helpful in avoiding this potential source of spurious correlation.

In some of the analysis below, we used information from the residential migration survey to identify individuals who had lived in the same area over the course of the exposure periods we were interested in, typically five to six years. Using summarized air pollution data from these periods, we tested hypotheses about the health status of individuals living in geographical proximity to these levels--rather than to pollution in some other geographical area.

2.2.4 Documentation and Cross Referencing of HIS Data

Appendix B to this report contains a listing of the variables used in our analysis of acute and chronic morbidity that originate in the HIS. This list cross-references the variable of interest, its location on the HIS Public Use Data Tape (available from the National Center for Health Statistics), and the corresponding question number on the HIS itself.

2.3 Air Pollution Data

2.3.1 SAROAD System

The pollution data used in the present study came from two sources. The first was the Environmental Protection Agency's Storage and Retrieval of Aerometric Data (SAROAD) system from which we obtained every hourly or 24-hour reading from every air pollution monitor in the U.S. for the year 1979. The second source was the National Aerometric Data Branch (NADB) time series data, also maintained by EPA. The following is a description of the SAROAD system as of 1979:

> In accordance with requirements of the Clean Air Act and U.S. Environmental Protection Agency (EPA) regulations for State Implementation Plans (SIPs), ambient air quality data resulting from air monitoring operations of state, local, and Federal networks must be reported each calendar quarter to the EPA. The EPA Storage and Retrieval of Aerometric Data (SAROAD) format is the established medium for transmittal of air data to EPA Regional Offices within 45 days after the end of each reporting period. EPA Regional Offices must, within an additional 30 days, forward data they have received to the EPA Aeormetric and Emissions Report System (AEROS), of which the SAROAD system is an operational party. AEROS is managed by the National Air Data Branch (NADB), Monitoring and Data Analysis Division, Office of Air Quality Planning and Standards.[10]

In 1979, the SAROAD files contained data on approximately 12,000 monitoring sites, 4,000 of which were then operational. The codes identifying each site are uniquely defined by five criteria:

(i) State Code
(ii) Area Code (city, or if not in city, county)
(iii) Site Code (001-999)
(iv) Agency Code
    A-E ... EPA Headquarters Groups
    F ... State Agency
    G ... County Agency
    H ... City Agency
    I ... District Agency
    J ... Private
    K ... Institution
    L ... Military
    M ... International Agency
    N ... Other Federal Nonmilitary Agency
    P ... EPA Regional Group
    Q ... World Meteorological Organization
    R ... World Health Organization
    Z ... Other
(v) Project Classification Code
    01 ... Population-oriented
    02 ... Source-oriented
    03 ... Background
    04 ... Complaint investigation
    05 ... Special studies
    06 ... Episode monitoring
    08 ... Global surveillance
    09 ... Duplicate sampling
    10 ... Continuous Air Monitoring Program

Thus, for example, a site identifier of 310660004F01 indicates that the monitoring site is located in New Jersey (31), specifically in Burlington County (0660), its number is 004, it is operated by a state agency (F), and the monitor is population-oriented (01). An identifier of 482890001F03 indicates a location in Virginia (48), specifically Shenandoah National Park in Page County (2890), at site 001, operated by a state agency (F), for background surveillance (03).

Other considerations are the type of sampler or monitor used in the data collection procedure and the frequency with which the data are collected (or, if continuous, are summarized). Insofar as collection methods are concerned, the number of different types of monitors varies with the pollutant in question. In 1979 suspended particulate monitors were all of the same type (HI-VOL Gravimetric), for example, whereas there were twelve types of sulfur dioxide monitors, eleven types of nitrogen dioxide monitors, and three kinds of carbon monoxide monitors. Two different monitoring methods were used for ozone: instrumental chemiluminescence and instrumental ultraviolet, with approximately twice as many monitoring sites using the former as the latter. Both methods produce data at one-hour frequencies, whereas for some of the other pollutants, the data frequency varies with the monitoring process being utilized.

2.3.2 Data Integrity and Completeness

Three basic statistics were used in our study to summarize individual pollution exposure over the two-week period for which we had acute health data and all were derived directly from the SAROAD raw data. These were the average of the fourteen highest daily one-hour readings for a two-week reference period; the highest single hourly reading over the reference period; and the average hourly concentration over the entire two-week period. Two measures of the completeness of the data at each monitor were also included in the air pollution data we matched to individuals. The first was the number of days during a two-week reference period for which there was at least one hourly or 24-hour reading (bounded between 0 and 14). The second was the total number of hourly observations available from a monitor during a reference week (336 maximum).

Tables 2-3 and 2-4 compare for two randomly chosen monitors the annual data derived from our method of summarizing pollutant concentrations by two-week reference periods. Our summary data compare exactly with the monitoring data summarized for each site in the NADB "Quick Look" Report (included following the monitoring data for each of the two sites). The annual summary data appear in the last row of Tables 2-3 and 2-4.

In addition to our concern about data completeness, we also analyzed the quality of the data in an effort to identify miscoded or otherwise inaccurate data. Because ozone is the focus of this study, the ozone data were subjected to the most detailed analysis. Nevertheless, all the air pollution data were screened using one or more outlier tests. Air quality data are usually screened by federal, state, or local reporting agencies prior to submittal of the data to the SAROAD system. Thus, it is not surprising that the ozone data exhibited reasonable consistency. Given more time the screening tests applied to the ozone data could be applied to any of the other

Table 2-3. EPA-SAROAD MONITOR PERFORMANCE BY HIS 1 REFERENCE DATES
SAROAD 07 330004F01 (GREENWICH, CT)

Reference Week	Date	Mean Daily Maximum	Days with at Least One Reading	Highest Hourly Reading	Mean Hourly Concentration	Hours > .0055	No. of Hourly Readings
1	JAN 1 - JAN 14	.0216	14	.0420	.0124	229	322
2	JAN 8 - JAN 21	.0242	14	.0420	.0166	211	268
3	JAN 15 - JAN 28	.0316	14	.0430	.0216	236	273
4	JAN 22 - FEB 4	.0389	14	.0530	.0268	321	327
5	JAN 29 - FEB 11	.0392	14	.0530	.0292	327	328
6	FEB 5 - FEB 18	.0406	14	.0480	.0304	328	330
7	FEB 12 - FEB 25	.0433	14	.0680	.0313	328	329
8	FEB 19 - MAR 4	.0383	14	.0680	.0235	293	315
9	FEB 26 - MAR 11	.0388	14	.0800	.0221	278	319
10	MAR 5 - MAR 18	.0497	12	.0800	.0349	255	275
11	MAR 12 - MAR 25	.0569	9	.0750	.0379	184	204
12	MAR 19 - APR 1	.0471	11	.0750	.0271	227	264
13	MAR 26 - APR 8	.0380	14	.0480	.0250	281	311
14	APR 2 - APR 15	.0402	14	.0500	.0286	290	306
15	APR 9 - APR 22	.0473	14	.0700	.0314	298	320
16	APR 16 - APR 29	.0570	14	.0820	.0336	295	325
17	APR 23 - MAY 6	.0661	14	.0860	.0391	312	331
18	APR 30 - MAY 13	.0760	14	.1250	.0408	314	331
19	MAY 7 - MAY 20	.0654	14	.1250	.0331	299	334
20	MAY 14 - MAY 27	.0564	14	.0950	.0303	294	333
21	MAY 21 - JUNE 3	.0782	14	.1400	.0361	293	334
22	MAY 28 - JUNE 10	.0938	14	.1760	.0429	295	330

Table 2-3 (continued)

Reference Week	Date	Mean Daily Maximum	Days with at Least One Reading	Highest Hourly Reading	Mean Hourly Concentration	Hours > .0055	No. of Hourly Readings
23	JUNE 4 - JUNE 17	.0927	14	.1050	.0464	306	328
24	JUNE 11 - JUNE 24	.0784	14	.1050	.0398	297	333
25	JUNE 18 - JULY 1	.0749	14	.1260	.0342	300	331
26	JUNE 25 - JULY 8	.0686	14	.1260	.0314	303	327
27	JULY 2 - JULY 15	.0808	14	.2040	.0344	283	330
28	JULY 9 - JULY 22	.1047	14	.2040	.0407	269	334
29	JULY 16 - JULY 29	.1134	14	.1680	.0418	277	330
30	JULY 23 - AUG 5	.1172	14	.1680	.0421	269	329
31	JULY 30 - AUG 12	.0910	14	.1550	.0392	289	333
32	AUG 6 - AUG 19	.0614	14	.1170	.0324	307	329
33	AUG 13 - AUG 26	.0669	14	.1180	.0325	294	331
34	AUG 20 - SEPT 2	.0792	14	.1180	.0324	271	336
35	AUG 27 - SEPT 9	.0657	14	.1100	.0275	280	335
36	SEPT 3 - SEPT 16	.0546	14	.0900	.0260	296	334
37	SEPT 10 - SEPT 23	.0518	14	.0900	.0245	269	334
38	SEPT 17 - SEPT 30	.0474	14	.0850	.0223	266	334
39	SEPT 24 - OCT 7	.0463	14	.0700	.0215	263	335
40	OCT 1 - OCT 14	.0382	14	.0700	.0182	210	336
41	OCT 8 - OCT 21	.0371	14	.0600	.0157	189	336
42	OCT 15 - OCT 28	.0432	14	.1050	.0171	222	336
43	OCT 22 - NOV 4	.0386	14	.1050	.0172	252	336
44	OCT 29 - NOV 11	.0293	14	.0400	.0139	207	336
45	NOV 5 - NOV 18	.0261	14	.0400	.0136	220	336

Table 2-3 (continued)

Reference Week	Date	Mean Daily Maximum	Days with at Least One Reading	Highest Hourly Reading	Mean Hourly Concentration	Hours > .0055	No. of Hourly Readings
46	NOV 12 - NOV 25	.0293	14	.0450	.0146	228	336
47	NOV 19 - DEC 2	.0279	14	.0450	.0143	192	336
48	NOV 26 - DEC 9	.0243	14	.0400	.0133	177	336
49	DEC 3 - DEC 16	.0199	14	.0300	.0096	139	336
50	DEC 10 - DEC 23	.0205	14	.0340	.0118	182	336
51	DEC 17 - DEC 30	.0242	14	.0340	.0161	233	336
Annual Statistics		.0538	51	.2040	.0270	6874	8432

OZONE (PARTS PER MILLION) CONNECTICUT 79-79

METHODS: HOURLY VALUES CHEMILUMINESCENCE-11, ULTRA VIOLET DASIBI CORPORATION-14, CHEMILUMINESCENCE RHODAMINE B DYE-15

SITE ID	LOCATION	COUNTY	ADDRESS	YR	NOBS	DAILY MAX 1-HR 1ST	2ND	3RD	VALS > .125 MEAS	EST	8HR VALID DAILY MAX	MISS DAYS ASS < STD	ME
070060123F01	BRIDGEPORT	FAIRFIELD CO	I...ETT & BARNUM	79	4318	.201	.195	.165	15	30.1	182	0	11
070175123F01	DANBURY	FAIRFIELD CO	TRAILER - OSBORN	79	3309	.182	.180	.170	14	37.0	138	0	11
070190123F01	DERBY	NEW HAVEN CO		79	3450	.233	.190	.182	14	35.9	142	1	11
070250123F01	ENFIELD	HARTFORD CO		79	3987	.165	.162	.146	11	24.0	167	1	11
070330004F01	GREENWICH	FAIRFIELD CO	BRUCE GOLF COURS	79	8432	.204	.185	.176	17	17.4	353	4	11
070350123F01	GROTON	NEW LONDON CO		79	1752	.186	.166	.163	6	29.9	73	1	11
070420123F01	HARTFORD	HARTFORD CO	HARTFORD TRAILER	79	8330	.235	.203	.178	15	15.5	350	3	11
070470001F03	LITCHFIELD CO	LITCHFIELD CO	MORRIS DAM	79	2458	.310	.230	.220	17	61.4	101	0	11
070700123F01	NEW HAVEN	NEW HAVEN CO		79	4113	.205	.177	.175	12	25.3	173	0	11

NATIONAL AEROMETRIC DATA BANK
QUICK LOOK REPORT

Table 2-4. EPA-SAROAD MONITOR PERFORMANCE BY HIB-1 REFERENCE DATES
SAROAD-148520007501 (Wood River, IL)

Reference Week	Date	Mean Daily Maximum	Days with at Least One Reading	Highest Hourly Reading	Mean Hourly Concentration	Hours > .0055	No. of Hourly Readings
1	JAN 1 - JAN 14	-9.0000	-9	-9.0000	-9.0000	-9	-9
2	JAN 8 - JAN 21	-9.0000	-9	-9.0000	-9.0000	-9	-9
3	JAN 15 - JAN 28	-9.0000	-9	-9.0000	-9.0000	-9	-9
4	JAN 22 - FEB 4	.0423	3	.0530	.0257	54	59
5	JAN 29 - FEB 11	.0442	10	.0690	.0246	203	226
6	FEB 5 - FEB 18	.0436	14	.0690	.0252	300	335
7	FEB 12 - FEB 25	.0389	14	.0590	.0233	294	336
8	FEB 19 - MAR 4	.0318	14	.0510	.0184	284	335
9	FEB 26 - MAR 11	.0299	14	.0370	.0178	283	329
10	MAR 5 - MAR 18	.0359	14	.0490	.0225	309	330
11	MAR 12 - MAR 25	.0362	14	.0490	.0230	328	336
12	MAR 19 - APR 1	.0329	14	.0400	.0204	317	336
13	MAR 26 - APR 8	.0361	14	.0500	.0235	320	334
14	APR 2 - APR 15	.0400	14	.0530	.0249	319	334
15	APR 9 - APR 22	.0464	14	.0730	.0259	312	336
16	APR 16 - APR 29	.0454	14	.0730	.0263	317	335
17	APR 23 - MAY 6	.0431	14	.0800	.0254	320	335
18	APR 30 - MAY 13	.0455	14	.0600	.0259	312	335
19	MAY 7 - MAY 20	.0547	14	.1070	.0288	299	329
20	MAY 14 - MAY 27	.0619	14	.1070	.0317	300	330
21	MAY 21 - JUNE 3	.0596	14	.0810	.0296	300	336
22	MAY 28 - JUNE 10	.0604	14	.0860	.0279	288	335

Table 2-4 (continued)

Reference Week	Date	Mean Daily Maximum	Days with at Least One Reading	Highest Hourly Reading	Mean Hourly Concentration	Hours > .0055	No. of Hourly Readings
23	JUNE 4 - JUNE 17	.0684	14	.0920	.0331	277	311
24	JUNE 11 - JUNE 24	.0703	14	.0920	.0360	274	294
25	JUNE 18 - JULY 1	.0604	14	.0860	.0289	277	318
26	JUNE 25 - JULY 8	.0626	13	.1070	.0320	276	312
27	JULY 2 - JULY 15	.0760	13	.1070	.0392	273	302
28	JULY 9 - JULY 22	.0756	14	.1060	.0398	302	326
29	JULY 16 - JULY 29	.0511	14	.0810	.0270	293	336
30	JULY 23 - AUG 5	.0526	14	.1050	.0204	257	335
31	JULY 30 - AUG 12	.0689	14	.1130	.0282	272	331
32	AUG 6 - AUG 19	.0661	14	.1130	.0296	298	332
33	AUG 13 - AUG 26	.0666	14	.1210	.0282	293	324
34	AUG 20 - SEPT 2	.0622	14	.1210	.0242	272	324
35	AUG 27 - SEPT 9	.0571	14	.0860	.0245	283	336
36	SEPT 3 - SEPT 16	.0646	14	.1310	.0276	281	336
37	SEPT 10 - SEPT 23	.0649	14	.1310	.0254	263	336
38	SEPT 17 - SEPT 30	.0672	14	.0900	.0259	253	336
39	SEPT 24 - OCT 7	.0594	14	.0900	.0233	251	336
40	OCT 1 - OCT 14	.0368	14	.0730	.0174	273	336
41	OCT 8 - OCT 21	.0299	14	.0630	.0157	273	336
42	OCT 15 - OCT 28	.0302	14	.0570	.0143	264	336
43	OCT 22 - NOV 4	.0281	14	.0570	.0130	249	313
44	OCT 29 - NOV 11	.0206	9	.0380	.0100	141	213
45	NOV 5 - NOV 18	.0288	5	.0570	.0137	58	105

Table 2-4 (continued)

Reference Week	Date	Mean Daily Maximum	Days with at Least One Reading	Highest Hourly Reading	Mean Hourly Concentration	Hours > .0055	No. of Hourly Readings
46	NOV 12 - NOV 25	.0309	10	.0570	.0158	170	225
47	NOV 19 - DEC 2	.0259	14	.0390	.0141	287	336
48	NOV 26 - DEC 9	.0249	14	.0430	.0123	256	332
49	DEC 3 - DEC 16	.0216	14	.0430	.0102	204	328
50	DEC 10 - DEC 23	.0184	14	.0300	.0090	193	332
51	DEC 17 - DEC 30	.0147	14	.0270	.0073	168	336
Annual Statistics		.0465	48	.1310	.0232	6391	7625

NATIONAL AEROMETRIC DATA BANK
QUICK LOOK REPORT

OZONE (PARTS PER MILLION) ILLINOIS 79-79

METHODS: HOURLY VALUES CHEMILUMINESCENCE-11, ULTRA VIOLET DASIBI CORPORATION-14, CHEMILUMINESCENCE RHODAMINE B DYE-15

SITE ID	LOCATION	COUNTY	ADDRESS	YR	#OBS	DAILY MAX 1-HR 1ST	2ND	3RD	VALS > .125 MEAS	EST	NBR VALID DAILY MAX	MISS DAYS ASS < STD	ME
147200007F01	SPRINGFIELD	SANGAMON CO	STATE FAIRGROUND	79	7046	.108	.087	.087	0	0.0	293	10	14
147550001F03	WATERLOO	MONROE CO	U.S. DEPT. OF AG	79	5912	.122	.088	.081	0	0.0	242	4	14
148920002F01	WAUKEGAN	LAKE CO	GOLF & JACKSON S	79	7541	.106	.158	.133	7	8.0	312	9	14
148220003F01	WHEATON	DU PAGE CO	111 N COUNTRY FA	79	6512	.112	.085	.082	0	0.0	272	2	14
148320007F03	WILL CO	WILL CO	SEWAGE PLANT	79	5010	.110	.105	.104	0	0.0	206	5	14
148320008501	WILL CO.	WILL CO	PARK DISTRICT OF	79	3700	.201	.109	.103	1	2.5	147	4	14
148600001F01	WINNEBAGO CO	WINNEBAGO CO	MAPLE ELEM. SCH.	79	3616	.084	.083	.083	0	0.0	152	1	14
148520007F01	WOOD RIVER	MADISON CO	WATER TRT PLT 54	79	7625	.131	.121	.114	1	1.1	314	8	14

pollutant data series used in this study by simply modifying the basic ozone quality control statistics. Based on the extensive analysis of ozone and screening of the other data, it appears that the air quality data are not subject to serious irregularities.

Ozone exhibits consistent diurnal and seasonal concentration patterns depending on a wide variety of meteorological factors. The tests selected to discover anomalies in the data reflect this pattern and may be grouped into tests designed to identify implausibly high concentrations based on historical data, and tests that basically evaluate the consistency of diurnal data.

In this study four tests were selected: the maximum hour and adjacent hour tests, and two variants of the spike test based upon percentage and absolute differences in adjacent diurnal hours. Table 2-5 presents the critical values used in the four tests. They reflect recognized seasonal and diurnal ozone concentrations, and check for an upper limit for particular ozone concentration. It is well known that ozone concentration varies by both season and day and while these critical values are somewhat subjective, the basic motivation in applying the tests is to identify data that warrant further examination.

The basic procedure for evaluation is relatively straightforward. The maximum hourly test simply searches for data exceeding a maximum upper bound. The second test compares an upper limit on the difference between adjacent hourly ozone readings. The third and fourth tests compare the middle of three consecutive hours in terms of absolute and

Table 2-5. Selected Quality Control Tests: Critical Values for SAROAD O_3 Hourly Values (Concentration in Parts per Million)

Data Stratification	Maximum Hour Test	Adjacent Hour Test	SPIKE[2] Test-1	SPIKE[3] Test-2
Summer-Day[1]	.501	.153	.102	300%
Summer-Night	.383	.102	.051	300%
Winter-Day	.255	.128	.102	300%
Winter-Night	.153	.102	.051	300%

[1] Reference weeks 18-44 are classified as "summer," while 1-17 and 45-52 are "winter."

[2] Under this spike test, the second of three consecutive hourly readings is flagged if it exceeds either endpoint by the stated value.

[3] Under this spike test, the second of three consecutive hourly readings is flagged if it exceeds either endpoint by 300%.

Reproduced from best available copy.

percent change in hourly ozone concentrations. Each of these tests are fully described elsewhere, as are specific results applying these tests to specific monitors.[11] The stratification of ozone data by season and time of day is given in column 1 of Table 2-5, whereas the critical values are reported in columns 2-5.

There are drawbacks to these tests. For example, it would be preferable if the critical values in Table 2-5 were available for separate geographic regions. Furthermore, the critical values themselves are subjective. Finally, there are relatively more sophisticated data screening tests available. However, it is very difficult and costly to obtain and evaluate air pollution data by regional grouping let alone by specific monitor. In our analysis critical values were only used to identify data that warrant further study. Finally, more sophisticated tests are themselves more costly and only recently are being applied to evaluate SAROAD data. Obviously, given more time and resources additional tests could be applied, but those tests conducted give us confidence about the relative consistency of the ozone data we used.

Approximately 660 SAROAD ozone monitors reporting daily data during 1979 were screened using the four tests described above. This includes all the air pollution data available to this study--approximately five million hourly observations for 1979. The number of ozone monitors in each state that had data flagged by the screening tests is presented in Table 2-6. We checked each hourly observation flagged by the screening tests. In every case the reading was a

Table 2-6. Number of Ozone Monitors Flagged in Data Screening Tests

State	Test		
	Maximum Hour	Adjacent Outlier	Spike
Arizona		1	1
Arkansas			1
California	8	34	51
Colorado			3
Connecticut		4	8
Delaware			1
District of Columbia			2
Florida		1	2
Georgia			2
Illinois	1	2	6
Indiana		2	4
Kentucky		2	5
Maryland			2
Massachusetts		3	10
Missouri	1	2	6
Montana		1	1
Nebraska			2
Nevada			1
New Jersey			1

Table 2-6 (continued)

State	Test Maximum Hour	 Adjacent Outlier	 Spike
New Mexico			1
New York		1	12
Ohio			8
Oklahoma			1
Pennsylvania	1	4	18

plausible one and apparently not the result of a malfunctioning monitor or an error in coding the data. The specific readings that were flagged are available upon request.

2.3.3 Multi-Year Averaged Data

As discussed above, several different measures have been used to summarize and verify the raw air pollution data for 1979 from EPA's SAROAD system. This was done to provide a view of the overall exposure to ozone and other air pollutants which individuals received in the different SMSAs around the United States. We have also summarized individuals' exposures to air pollution over a number of years using long-term data from the air pollution monitors assigned to them.

The summary data used in this exercise are those in the National Aerometric Data Base, Yearly Frequency Distribution. The data tape made available to us included annual summary measures for ozone, suspended particulates, nitrogen dioxide, carbon monoxide, sulfur dioxide, and lead, for the years 1974 through 1980. No sulfate data were available. For each pollutant, the method of collection, number of annual observations, arithmetic and geometric means, and highest and second highest annual values were recorded. It was also reported whether the annual summary measure reported met the EPA criteria for data completeness.

In attempting to compile multi-year summary measures of pollution to characterize individuals' long-term exposure to ambient levels, we have used only the data meeting EPA completeness criteria for the years

1974 through 1979. We matched on a monitor-by-monitor basis such data to each individual's other recorded data in a method completely analogous to the matching procedure described above. It will be recalled that the monitor matching is based on monitors in operation in 1979, and that data were deleted for all monitors more than twenty miles from the centroid of the census tract in which a particular individual was living. For each of the six years 1974-1979, the summary data were averaged over all monitors within twenty miles. Finally, these six annual averages were simply averaged to obtain the summary measures of interest. In each of the averaging processes, missing data were ignored and averages were calculated based on the existing data. We recognize that our method obscures the variance about the mean at an individual monitor over the course of one year, between all the monitors within twenty miles for a given year, and over the multi-year "grand averaging." In subsequent work we hope to include measures of dispersion as well as central tendency.

An example is presented to clarify this procedure. Table 2-7 contains illustrative annual average hourly ozone data for one individual. The first row shows the distances in miles to the ten nearest ozone monitors for the individual. Rows two through seven are the 1974 through 1979 concentrations at these monitors for data which meet EPA completeness criteria. Missing values are indicated by dots. The final column in rows two through seven presents the six annual averages, subject to the twenty-mile cutoff. The figure in the lower right hand corner of the table shows the value derived as this

Table 2-7. Summary of Long-term Annual Average Hourly Ozone Concentration for a Representative Individual

Annual Average Ozone Concentration

Monitor no.	1	2	3	4	5	6	7	8	9	10	Average of monitors 1-4[1]
Distance	8.9	13.2	14.0	16.2	21.9	22.8	22.9	23.6	28.0	32.0	
Year											
1974	0.19	.	.	.	.	.	.	.	.015	.	.019
1975	.	.	.	.	.	.	.	.	.012	.	.
1976	.023	.	.	.	.	.	.		.011		.023
1977	.020	.	.	.	.	.	.	.018	.012	.	.020
1978	.019	.024	.013	.	.	.	.	.013	.014	.022	.0187
1979	.	.	.014	.	.	.018	.029	.015	.	.	.014
										Overall average:	.0189

---

[1] Only monitors 1-4 are used to compute the within-year and multi-year averages in this example because monitors 5-10 are more than twenty miles from the individual's home. This is the procedure used in the creation of the multi-year data.

individual's long-term annual arithmetic average ozone concentration (.0189 in this case).

The summary measures constructed from the NADB Yearly Frequency Data are as follows. For ozone, sulfur dioxide, carbon monoxide, and nitrogen dioxide, the six-year averages of both the arithmetic means and the second yearly high values were computed. For suspended particulates, the average of the geometric means as well as the second yearly high values were computed. Also calculated for each individual were the number of monitor values used in calculating each yearly average as well as the number of years used in calculating the overall average.

When contrasting the multi-year annual averages with the annual averages calculated from the 1979 data alone, one expects the summary to differ. For example, for the 14,441 adults in our sample Table 2-8 shows the simple correlation matrix of the five different long-term measures of ozone used in the present study. The statistics calculated from the 1979 hourly data are also subject to the twenty-mile cutoff rule, i.e., readings from monitors at a distance of twenty miles or greater from the individual's census tract are assigned a missing value. The ozone measures are indexed as follows:

O3(1) = Average daily 1 hr. maximum, 1979

O3(2) = Maximum hourly reading, 1979

O3(3) = Average hourly reading, 1979

O3(4) = Annual arithmetic average, 1974-79

O3(5) = Average of second-high yearly reading, 1974-79

O3(i), i=1,2,3, are derived from the raw hourly data for 1979 while O3(j), j=4,5 are derived from the NADB Yearly Frequency data using the methods described in the above paragraphs. The correlation matrix is as follows:

Table 2-8. Simple Correlation Coefficients Between Alternative Measures of Long-term Exposure to Ozone

	O3(1)	O3(2)	O3(3)	O3(4)	O3(5)
O3(1)	1.00	.783	.886	.700	.697
O3(2)		1.00	.549	.469	.789
O3(3)			1.00	.664	.451
O3(4)				1.00	.593
O3(5)					1.00

A number of important assumptions were made in compiling the multi-year average pollution concentrations (for 1974-79). When calculating the 1974-79 annual summary statistics, different years were given the same weight even though there were often more readings available for a particular monitor in one year than another. Also, the differences in the number of observations across all the monitors within 20 miles were not weighted in compiling summary statistics for any one year. Differences across years in the quality of monitoring itself have not been accounted for here. These problems combine to make it difficult to derive any single number that characterizes an individual´s six-year exposure to air pollution. Of course, any single measure of long-term exposure is bound to be imperfect. However, we believe the approach taken here--which uses as many as six years of air pollution data--is a step in the right direction, one which we have not seen elsewhere. Because the chronic effects of air pollution on individuals´ health are an area of concern, further efforts to measure long-term exposure to pollution are important. As the quality of monitoring data continues to improve, the payoffs from such a task will be large.

2.3.4 Description of the Pollution Variables

A large number of different air pollution measures or variables were used in this report so it may be helpful to describe in some detail the system we used to name these variables. The basic descriptors for the pollution variables are the prefixes:

O3 -- Ozone
SP -- Total suspended particulates
S2 -- Sulfur dioxide
S4 -- Sulfates
N2 -- Nitrogen dioxide
CO -- Carbon monoxide

These descriptors are always the first two characters in the names of the pollution variables. Thus, any variable beginning with SP is some measure of total suspended particulates. Similarly, the prefix N2 signifies some measure of nitrogen dioxide. To indicate whether the appropriate pollution measure comes from each individual's first nearest, second nearest, ..., tenth nearest monitor for each type of pollutant, the third and fourth characters in the variable name are indexed from 01 to 10. The fifth character in the variable name is either R, L, M, or A. Respectively, this signifies data specific to the two-week reference period; the two-week period immediately preceeding the reference period; the average of these adjacent two-week periods (weighted by the number of observations available from each); or the entire year. Combined, the sixth and seventh characters range from 01 to 06. These six suffixes describe the particular summary measures for each pollutant, at each monitor, and over each possible summary period, as follows:

01	Mean of daily maximum for summary period
02	Number of days for which at least one daily observation was available
03	Highest hourly or daily reading
04	Mean hourly or daily concentration
05	Number of hours or days in a summary period in which reading exceeded some specified value
06	Number of hourly or daily observations available during the summary period

where hours/days refers to the maximum frequency of data availability.

Therefore, for example, 0301R01 signifies the average daily one-hour maximum of ozone during the two-week reference period at the monitor nearest each individual's home; N207L04 is mean hourly concentration of nitrogen dioxide during the two-week period preceeding the two-week reference period at the seventh nearest monitor to the individual's home; and SP04A05 is the highest total suspended particulate reading recorded at the fourth nearest monitor.

Measures where the third character is "N" are those measures obtained from the monitor nearest the individual for which data were available during the summary period. In such cases, one of R, L, M, or A is the fourth character, with the same meaning as above, and 01, ..., 06 again index the type of summary measure, but are now located in the fifth and sixth character positions of the variable names. For example, 03NR01 is the average daily one-hour maximum ozone concentration for the individual's two-week reference period from the nearest monitor for which such data exist, while 03NL04 is the average hourly concentration of ozone for the two-week period preceeding the individual's two-week reference period from the nearest monitor for which such data are available.

Not listed above are the variables summarized from the NADB time series data described in the preceeding section. We identify these variables in the following way. Pollutants are again indexed by 03, SP, S2, CO, or N2 (sulfate data were not available for 1974-1979 from the NADB system tape we used). These identifiers occupy the first two character positions of the variable names. "AN" in the third and fourth character positions means that the variable is a six-year annual summary derived using the methodology described in Chapter 2. "GEAV" and "ARAV" in the fifth through eighth character positions mean geometric mean and arithmetic mean, respectively. "MAX2" in character positions five through eight indicates the second highest reading for the year, averaged in the manner described above. "GENU," "ARNU," and "M2NU" represent the number of annual summary data points over which the geometric mean, arithmetic mean, and second maximum averages were calculated.

For example, O3ANARAV is the six-year average of annual arithmetic average hourly ozone concentrations; if O3ANARNU=5, the average was calculated using data from five years. Variables prefixed by O3, SP, S2, N2, or CO, with the third and fourth characters indexed over 74, ..., 79 are the measures of data completeness. In such measures, the fifth through eighth character positions in the name are "NUGM," "NUAM," or "NUM2," signifying, respectively the number of monitor values satisfying EPA and our data quality criteria (described above) for geometric means, arithmetic means, and second annual maxima. Thus, for example, SP77NUGM=3 means that there were data from three monitors satisfying completeness and proximity criteria for the geometric mean of suspended particulates in 1977. Note that geometric averages were used exclusively for suspended particulates, while arithmetic averages were used only for ozone, sulfur dioxide, carbon monoxide, and nitrogen dioxide.

We recognize the complexity of our system for naming air pollution variables. It is necessitated by our desire to experiment with a number of different measures of exposure to the various pollutants. If it is any consolation to the discouraged reader, most of the analysis below uses only one or two measures of each air pollutant. In the analysis of acute morbidity in Chapter 4, we typically used the nearest monitor for which "complete" data were available. Thus, the variables almost always look like O3NR01, ..., CONR01. In the analysis of chronic morbidity, we used air pollution data averaged over all of 1979, or multi-year averaged data. Thus, the pollution measures look like O3NA01, ..., CONA01 (for the 1979 data) or O3ANARAV, ..., COANARAV. In actuality, it is not overly difficult to identify the air pollution variables being used.

2.3.5 Matching Individuals to Monitors

The algorithm used to match air pollution data to the individuals in the Health Interview Survey and in our final data set used the census tract in which individuals reside as the focal point. It takes a two-step approach to the matching process. In the first step, the geographic coordinates of the centroids of each census tract are calculated using the concept of oriented polygons. This procedure allows for the fact that miles per degree longitude depends on latitude. The data used in this step are from the Census Bureau's Urban Atlas Files. Importantly, they are only available for census tracts within Standard Metropolitan Statistical Areas (SMSAs). This meant that the matching procedure could only be used for residents of SMSAs. The implications of this restriction for possible sample selection bias are discussed in a later section of this chapter.

Once the geographical coordinates of the centroids were determined, the second step was a relatively straightforward one. It involved "sweeping out" from the centroid of a census tract and identifying the nearest ten pollution monitors for each of ozone, suspended particulates, sulfur dioxide, nitrogen dioxide, sulfate ion, carbon monoxide, total oxides of nitrogen, and lead. This step made use of EPA data on the geographic coordinates of all air pollution monitors, contained in their Directory of Air Quality Monitoring Sites. Using this sweeping algorithm and the unique identifiers of each monitoring site described above, the distances in miles from census tract centroid to each monitor were established for each qualifying individual in the dataset. Having thus matched individuals to monitors, it was then straightforward to assign them the appropriate air pollution data from the SAROAD system.

Throughout the rest of this report we refer to the monitors "nearest individual's homes." We do so only for convenience of exposition, however; this discussion of matching should make it clear that all individuals are assigned data from monitors nearest (then second nearest, and so on) the centroid of the census tract in which they lived at the time of the 1979 HIS.

2.4 Meteorological Data

The meteorological data used in the following chapters derived from data collected and published by the National Climatic Center, National

Oceanic and Atmospheric Administration (NOAA), United States Department of Commerce. All data used were specific to 1979.

Two NOAA data bases were used to generate data for our analysis: monthly and annual values from the Local Climatological Data--Annual Summaries for 1979; and daily values from the WBAN Summary of Day, Deck 345, for 1979. From the former, monthly and annual values summarized from data recorded at the weather stations that constitute the NOAA monitoring network were available for average temperature, relative humidity for the hour closest to noon, average windspeed, precipitation in inches, snowfall, heating degree days, and cooling degree days. From the latter data, the daily values of interest for purpose of this study were maximum and minimum 24-hour temperature, 24-hour precipitation, 24-hour snowfall, 24-hour peak instantaneous wind speed, and presence or absence during the 24-hours peak instantaneous wind speed, and presence or absence during the 24-hours period of the following: smoke/haze, rain, or snow; and percent of possible sunshine. All told, this gave sixteen measures of climatological influences at the weather station level, although the overlaps (e.g., monthly versus daily precipitation readings) are obvious.

Due to the nature of the questions regarding health status in the Health Interview Survey, data recorded in monthly and daily frequencies were inappropriate to characterize individuals' climatological exposure. As explained above in the detailed description of the Health Interview Survey, individual responses were specific to two-week, six-month, or annual recall periods. Annual data are suitable to characterize climatological exposure when one is exploring the possible determinants of chronic morbidity. However, when examining acute morbidity during a particular two-week period (as collected in the HIS) neither monthly nor daily data are ideally suited to the task. Rather, data measured at daily and monthly frequencies must be transformed into measures representative of exposure to heat, precipitation, etc., during the two-week recall period specific to an individual's Health Interview Survey date. The description of the methodology used to perform this task follows.

For the monthly data, a simple weighted-averaging process was used. The weights were the number of days in the individual's two-week recall period falling in the month or months spanned by the two-week recall period divided by fourteen. For example, if the variable of interest is X and the recall period was March 31 - April 13, then the average, Y, was calculated as $Y = [(1/14)*X_{Mar} + (13/14)*X_{Apr}]$. For recall periods falling entirely within one calendar month, the monthly value of the variable was used as the two-week proxy. Variables calculated with this averaging method were suffixed by "RF", e.g., PRECIPRF represents the averaged monthly precipitation. Annual data were suffixed by "AN".

Data available on a daily basis were aggregated to form two-week measures. Four prefixes were used to describe the statistics calculated from the daily climatological data: "NU" signifies number of days over the two week period for which the relevant data were available; "MX" signifies the maximum recorded value during the two-week period for the variable in question; "MN" signifies the minimum recorded value; and "AV" signifies the mean calculated over all available observations. An example should serve to clarify things. If the daily values for maximum temperature recorded during the individual's two-week recall period were (75,80,85,n/a,n/a,85,80,75,n/a,n/a,80,80,80,n/a), then the following variables would assume the values listed:

NUMAXTMP...........9

MXMAXTMP..........85

MNMAXTMP..........75

AVMAXTMP..........80

Obviously, not all of the four statistics described above will be applicable in all cases.

The variables derived from the daily meteorological data and their descriptions are as follows:

Daily Maximum Temperature

MXMAXTMP, AVMAXTMP, NUMAXTMP

Daily Minimum Temperature

MNMINTMP, AVMINTMP, NUMINTMP

Daily Peak Gust

MXPKGUST, AVPKGUST, NUPKGUST

Daily Presence/Absence of Rain (1,0)

AVRAINYN, NURAINYN

Daily Presence/Absence of Snow (1,0)

AVSNOWYN, NUSNOWYN

Daily Presence/Absence of Smoke or Haze (1,0)

AVSMOKHZ, NUSMOKHZ

Daily Precipitation in Inches

MXPRECIP, AVPRECIP, NUPRECIP

Daily Snowfall in Inches

MXSNOWFL, AVSNOWFL, NUSNOWFL

Daily Percent of Possible Sunshine

MXPCSUNP, AVPCSUNP, NUPCSUNP

The process for matching weather stations to individuals was based on the Standard Metropolitan Statistical Area (SMSA) in which the individual resides. For each SMSA in the data set, one and only one weather station was assigned. In most cases, only one weather station was located in or near the SMSA and in these cases, the data recorded at that station were assigned to all residents of the SMSA. In a few cases, more than one station was available (e.g., for New York City, there were data from Central Park, LaGuardia Airport, and JFK Airport), in which case individuals were assigned data from the nearest weather station.

2.5 Other Data

In addition to the data described above pertaining to individuals' health and socioeconomic status and their exposure to air pollution and meteorological factors, there are several other types of data that deserve mention here. Certain of these data describe potential determinants of health that have been included in previous epidemiological investigations, although not to our knowledge in a study of this size and scope. Other data measure variables that have not before been used to analyze acute and chronic morbidity.

2.5.1 Pollen

In order to isolate the health effects associated with ozone and other pollutants, it is necessary to control for as many other influential factors as possible. The exclusion of important explanatory variables leads to biased estimates of the parameters on the included variables, unless all (vector) elements of the set of excluded variables are orthogonal to all (vector) elements of the set of included variables. In any event, the constant term, if included, will be biased.

The Health Interview Survey and the air pollution data from the Environmental Protection Agency's SAROAD system, along with other relevant data, provide information about many of the factors that might affect acute and chronic morbidity. However, none of these sources provides data on one potentially important determinant of health status--airborne concentration of pollens in the geographic areas of interest. Since observed ill health--particularly respiratory discomfort--may be due not to air pollutants, but rather to airborne allergens, it is clear that controlling for such influences is important if reliable parameter estimates are desired.

Unfortunately, pollen data are neither widely available nor highly reliable. To our knowledge, there exist no federal government sources. After investigating a number of possible sources, we settled on data amassed by the Pollen and Mold Committee of the American Academy of Allergy (AAA). In its annual report, AAA publishes data on a broad array of pollens monitored at various sites nationwide.

There are several problems with these data. First, not all AAA sampling stations collect data on all kinds of pollens. Some stations collect only ragweed data, while others monitor up to forty different pollen types. Second, different stations employ different--and noncomparable--types of collection devices. The most common of these are the Durham Sampler (DS) and the Intermittent Rotorod (IR), although others exist. Some of the methods use planar measures per unit time (e.g., particles/cm^2/24 hours), while others use volumetric measures (e.g., particles/m^3). There are no generally acceptable methods of conversion between these two measurement types, not even at a "rule-of-thumb" level. This limits the number of sites for which comparable data are available. Third, not all stations collect data every day of each month, i.e., figures are presented monthly for aggregate particles collected and total collection days. The best that can be done here--absent other information--is to assume that the sampling is random, that the average daily counts for the sample days in fact represent the overall averages.

However, we decided the AAA data could be useful in our efforts. Twenty-five stations in the metropolitan areas surveyed by NCHS collect data using the DS method. This guaranteed a large number of individual observations while simultaneously allowing the data to be derived from a single sampling method, thus ensuring comparability. We used DS data in our statistical work.

Although we would prefer to have data on all pollens in each SMSA, several medical experts suggested that a ragweed-only index would be an acceptable measure of pollen-related maladies, indeed by far the best single measure.[12]

In the analysis below, we test the seriousness the data limitations we face by estimating models with a subsample of individuals for whom there exist pollen data and comparing the estimated coefficients with those from regressions where no pollen data are included.

2.5.2 Indoor Air Pollution

There is a growing awareness among environmental health scientists of the risks posed by indoor air pollution.[13] Concentration gradients exist for a variety of air pollutants between indoor (residences, for example) and ambient air. Carbon monoxide and suspended particulates are likely to range higher in homes occupied by cigarette smokers; radon gas, which is trapped indoors, is likely to be high in certain sections of the country; nitrogen dioxide is likely to be higher in homes that use gas stoves for cooking; while ozone, sulfur dioxide, and sulfate particles are likely to be higher in ambient air.

The magnitude of the gradient reflects the strength of the pollutant sources and the ventilatory capacity of the indoor environment; increased ventilation acts to reduce the gradient. In buildings with potent sources--whether inefficient space heaters, gas ovens, or smokers--and low rates of ventilation (tightly sealed buildings), peak indoor concentrations can exceed those experienced outdoors, and even exceed the NAAQSs. Thus, one would like information about such indoor exposures in order to isolate health effects that might arise from exposures to outdoor concentrations. Unfortunately, no information about cooking fuels is elicited in the HIS, nor is any other information that could provide some guide to indoor exposures to air pollution, apart from the information in the smoking supplement.

Although no individual data are available, it is possible to draw some rough inferences about likely indoor exposures to carbon monoxide and nitrogen dioxide from natural gas for the individuals in our sample. Using the Provisional Estimates of Social, Economic, and Housing Characteristics of the 1980 Census of Population and Housing, data were collected on the number of occupied housing units in each of the thirty-eight SMSAs in the United States with greater than one million persons that use natural gas (either piped or bottled) as the primary cooking fuel. We converted these to percentages of occupied housing units and, in some of the regressions below, assigned all households in a particular SMSA the average for that SMSA. This ranged from a low of 4 percent in the Seattle-Everett SMSA to a high of 88 percent in the New York-New Jersey SMSA. The figure can be interpreted as the probability that the household cooked with natural gas, and hence the probability that its members were exposed to possibly high levels of carbon monoxide and nitrogen dioxide indoors. Individual data comparable to those on health and socioeconomic status would be preferred. However, the aggregate data from the census permitted a crude control for indoor exposures.

2.5.3. Paid Sick Leave

One of the major interests in this report is the relationship between air pollution and time lost from work, other factors held constant. One of these other factors which may influence work loss--and hence which should be controlled--is the availability and amount of paid sick leave that a worker enjoys. Two individuals may be identical in every respect studied here but for the amount of paid sick leave they are granted each year; if so, more work loss might be expected from the individual with more paid sick leave because a day lost from work will not cost him (or her) as much as it will cost the other worker (assuming that the limits are binding). The presence of paid sick leave may therefore lower the "threshold" of illness that must be reached before a worker stays home from the job.

An alternative hypothesis might be advanced, however. Paid sick leave may enable workers to stay home and rest during the early stages of an illness, thus reducing their chances of more serious illness that could lead to prolonged work loss. If so, workers with more paid sick leave might be expected to miss less work each year, ceteris paribus, than workers with less. At any rate, paid sick leave should play some role in explaining the amount of work loss.

No questions were asked on the 1979 HIS about paid sick leave. However, in 1974 a supplement on conditions of current employment was included in the HIS in which each working respondent was asked how many days of "wage loss coverage (sick leave, etc.)..." they received annually.[14] We averaged the reported paid sick leave for 1974 by occupation and industrial category. Then, assuming that the combined industry-occupational averages have remained the same since 1974, we assigned the same number of paid sick leave days to individuals in various industries and occupations for 1979. While it would have been preferable if the 1979 HIS had contained a question about paid sick leave, using the 1974 data from the same survey seemed to be a reasonable compromise. It is unlikely that the number of days of paid sick leave will have changed much since 1974, and even more unlikely that the relative rankings between industries and/or occupations will have changed.

2.5.4 The Pollutant Standard Index

Individuals must venture outside during polluted periods if one is to observe a relationship between either acute or chronic morbidity and air pollution. Yet it is possible that if individuals are made aware of periods of poor air quality, they would forego or reduce outdoor activities during these periods. If such behavior were widespread, the observed relationship between acute morbidity and air pollution might be negative, rather than as expected.

Conversely, information or perceptions about current levels of air quality might have a different effect. Individuals aware that the air quality during a particular period was very poor might become even more aware than usual of respiratory and other ailments and show an expectational effect. Since acute and chronic morbidity is self-diagnosed in the Health Interview Survey, this possibility is pertinent to the present study.

To explore the possible role played by information about air quality conditions, we obtained from the Environmental Protection Agency a summary of the metropolitan areas in the United States in which air quality conditions were regularly reported in 1979. In particular, those areas reporting the Pollutant Standard Index (PSI)--a convenient summary measure of air quality--were identified.[15] Adoption of the PSI was used to proxy the availability of ambient pollution information. A dummy variable was created for each area taking on the value 1 if the PSI was reported in 1979, and the value 0 if not.

Quoting from an EPA report:

> The PSI is presently being reported in 94 urbanized areas, accounting for 89 million people. Of these 94 urban areas, 39 areas have populations greater than or equalling 500,000; 39 have populations ranging from 200,000 to 500,000; and 16 have populations less than 200,000.

The PSI adoption dummy, PSIY1NO, was matched to the adults data set by SMSA identifier. In the sample of 14,441 adults, 12,615 (87.4%) lived in areas where the PSI was reported while 1,826 (12.6%) lived in areas where it was not. The EPA report used by RFF to construct this variable is available upon request.

2.5.5 Availability of Medical Care

In addition to the determinants discussed above, acute and chronic morbidity may depend in part on access to medical care. Those who can receive medical attention quickly and inexpensively can be expected to enjoy better health than those for whom medical care is an expensive and time-consuming activity. Unfortunately, the HIS contains no information about individual access to medical care. Rather, it concerns itself with acute and chronic illness, physician visits and so forth, without indicating the lengths to which an individual had to go to obtain that care.

In an effort to get some idea of its possible effect on acute and chronic morbidity, we included in our data base two measures of access to medical care: the per capita number of physicians in the SMSA in which an individual lives; and the per capita number of hospital beds in that SMSA. These data were derived from the U.S. Department of Commerce, National Technical Information Service (NTIS), Bureau of Health Manpower Area Resource File (ARF). The physicians and hospital data were derived, respectively, from the health manpower and health facilities sections of this file, while SMSA population was drawn from the population section of the same file.[16]

2.6 Problems of Sample Selection

When using data derived from other than purely randomized sampling procedures, one must understand the implications of and be prepared to correct for sample selection bias. This is most important when one goal of research is to extrapolate results from estimation based on samples to a group not represented in the observations over which the parameters were estimated. The problem has been summarized as follows:

> Sample selection bias may arise in practice for two reasons. First, there may be self selection by the individuals or data units being investigated. Second, sample selection decisions by analysts or data processors operate in much the same fashion as self selection.
>
> There are many examples of self selection bias. One observes market wages for working women whose market wage exceeds their home wage at zero hours of work. Similarly, one observes wages for union members who found their nonunion alternative less desirable. The wages of migrants do not, in general, afford a reliable estimate of what nonmigrants would have earned had they migrated. The earnings of manpower trainees do not estimate the earnings that nontrainees would have earned had they opted to become trainees. In each of these examples, wage or earnings functions estimated on selected samples do not, in general, estimate population (i.e., random sample) wage functions. Comparisons of the wages of migrants with the wages of nonmigrants (or trainee earnings with nontrainee earnings, etc.) result in a biased estimate of the effect of a **random** "treatment" of migration, manpower training, or unionism.
>
> Data may also be nonrandomly selected because of decisions taken by data analysts. In studies of panel data, it is common to use "intact" observations. For example, stability of the family unit is often imposed as a requirement for entry into a sample for analysis. In studies of life cycle fertility and manpower training experiments,

> it is common practice to analyze observations followed for the full length of the sample, i.e., to drop attriters from the analysis. Such procedures have the same effect on the structural estimates as self selection: fitted regression functions confound the behavioral parameters of interest with parameters of the function determining the probability of entrance into the sample.[17]

This researcher then summarized the problem: "The critical question is 'why are the data missing?'"

In our case, the discussion above has demonstrated "why the data are missing." The obvious follow-up is how to deal with the missing data problem.

In our study if an adult was given a supplemental questionnaire about smoking, he was eligible for inclusion in our final sample if we could match him to at least one air pollution monitor of any kind. But this required that the individual live in a Standard Metropolitan Statistical Area (SMSA). Not all residents living within SMSAs are included in our final sample, but no one living outside SMSAs is included. The following describes our working hypotheses for the sampling aspects of this project.

Let

S = final sample

M = set of individuals receiving the smoking survey

C = the set of children (aged 16 and under)

P = the set to whom monitors could be matched

V = the set of individuals surveyed living in SMSAs

R = the set of individuals surveyed not in SMSAs

T = the overall sample ($\#T = 110{,}530$)

$C_Q(A)$ = the complement of Q in the universe A

We then have

$$S = \{ i \mid i \in (C \cap P) \cup (C_C(T) \cap P \cap M)\} \qquad (1)$$

$$\Pr(i \in S \mid i \in R) = 0 \qquad (2)$$

$$\Pr(i \in V \mid i \in C \cup (C_C(T) \cap M)) = f(X_{1i};\beta_1) \qquad (3)$$

$$\Pr(i \in S \mid i \in V) = \Pr(i \in S \mid i \in T) - \Pr(i \in S \mid i \in R) = g(X_{2i};\beta_2) \qquad (4)$$

Identity (1) is merely definitional; (2) states that S contains no individual living outside an SMSA; (3) states that the probability that an individual resides in an SMSA given that the individual is in the entire sample (and given that a smoking survey was administered, if adult) depends on a set of characteristics of the individual, X_{1i}, a parameter vector, β_1, and a functional relationship f( ); (4) states that the probability than an individual is in the final sample given that the individual resides in the SMSA depends on a (perhaps different) set of individual characteristics, X_{2i}, a parameter vector, β_2, and a functional relationship g( ).

The first question to explore, then, relates to (3). What factors determine whether or not an individual resides in an SMSA? We want to consider

$f: X_{1i} \rightarrow [\ 0,\ 1\ ]$, such that

$$\Pr(\ i \in V \mid i \in C \ \cup (C_C(T) \cap M)) = f(X_{1i};\ \beta_1) \qquad (5)$$

We will now explore a somewhat restricted form of this hypothesis:

$$\Pr(\ i \in V \mid i \in C_C(T) \cap M) = f(X_{1i};\ \beta_1) \qquad (6)$$

Our approach is as follows. We assign to each adult receiving the smoking survey (i.e. all $i \in M$) a binary (0,1) variable called SMSA defined as

$$\begin{aligned} \text{SMSA} &= 1 \ \text{ iff } i \in V \qquad (7) \\ &= 0 \ \text{ iff } i \in R \end{aligned}$$

We then consider a model of the form

$$\Pr(\text{SMSA}_i = 1) = F(X_{1i}\ \beta_1) \qquad (8)$$

where F( ) is a distribution function. Our task is to estimate the parameter vector β_1, given X. We will assume that F( ) follows a logistic distribution, so that

$$\Pr(\text{SMSA}_i = 1) = (1 + \exp(-X_{1i}\ \beta_1))^{-1} \qquad (9)$$

The X_1 vector has the following elements

NE.........Northeast region dummy
NC.........Northcentral region dummy
SO.........South region dummy
SEX........Sex, 1 if male, 0 if female
AGE........Age in years
RACE.......Race dummy, 1 if white, 0 else
MAR........Marital status, 1 if married, 0 else
EDUC.......Years of schooling
INC........Pseudo-continuous measure of income, in $
CONS.......Vector of unity elements

The $\hat{\beta}_1$ vector estimated by maximum likelihood estimation of (9) is as follows, with asymptotic t-scores in parentheses.

NE	.115 (2.25)
NC	-.607 (13.20)
SO	-.977 (22.20)
SEX	-.057 (1.90)
AGE	.0027 (3.14)
RACE	-1.033 (18.78)
MAR	-.453 (13.73)
EDUC	.043 (8.60)
INC	.0000385 (22.65)
CONS	1.257 (13.23)

All estimated parameters are significant at above 95% except for the parameter on SEX, with most parameters significant at well above 99%. For example, these results indicate that people living in the northeast part of the country are more likely than those in the west (the excluded dummy) to live in SMSAs. Similarly, SMSA residents are more likely to be elderly, nonwhite, unmarried, more educated, and have higher incomes than those outside SMSAs.

A second question pertains to (4) above. That is: given that an individual resides in an SMSA, what are the factors influencing whether or not such an individual will be included in our final sample? We want to consider, given an attribute vector X_{21} and parameter vector β_2,

$$g\colon X_{21} \longrightarrow [\,0,1\,], \text{ such that}$$

$$\Pr(\, i \in S \mid i \in V\,) = g(\, X_{21};\ \beta_2) \qquad (10)$$

Again, we look at a restricted version of this,

$$\Pr(\, i \in S \mid i \in V \cap C_C(T) \cap M\,) = g(\, X_{21};\ \beta_2) \qquad (11)$$

Now we consider the possibility that -- in addition to the elements of X_1 -- the residence of an individual within or outside a central city of the SMSA will influence the probability of being included in the final sample. Thus, X_2 equals X_1 plus the additional element of INCITY, where INCITY = 1 if the individual resides in central city areas of SMSAs and = 0 if the individual resides outside central city areas of SMSAs but still within SMSA boundaries. The binary dependent variable is now called INSAMP, defined as

$$\begin{aligned} \text{INSAMP} &= 1 \text{ iff } i \in S \\ &= 0 \text{ iff } i \in C_S(V) \qquad (12) \end{aligned}$$

Again we posit a logit specification, so that

$$\Pr(\text{INSAMP}_i = 1) = (1 + \exp(-X_{2i}\,\beta_2))^{-1} \qquad (13)$$

With asymptotic t-scores in parentheses, the maximum likelihood estimator of $\hat{\beta}_2$ yields,

NE	.790	(13.33)
NC	.680	(11.72)
SO	.050	(0.940)
INCITY	.637	(14.29)
SEX	.109	(2.67)
AGE	.015	(11.19)
RACE	-.054	(0.819)
MAR	-.270	(5.82)
EDUC	-.068	(9.30)
INC	-.0000138	(5.76)
CONS	1.450	(11.33)

Only the parameters estimated on DUMSO and RACE fail to attain significance at the 99% level. Clearly, the INCITY effects are overwhelmingly significant, indicating that individuals residing in the central city areas of SMSAs are more likely to be selected into the final sample--other things equal--than those residing outside such areas. This follows from the tendency for pollution monitors to be more densely bunched in central city areas than in outlying areas within SMSAs.

(Because the question of the determinants of Pr(INCITY=1) is itself relevant, we estimated a third logit model which indicated that RACE and INC, both negatively, were the most significant explanatory variables as judged by asymptotic t-scores. The full results of this estimation are available on request.)

We have in effect analyzed two conditional probability models, first, the probability of living in an SMSA given that a smoking survey was administered, and second, the probability of being in the final sample given SMSA residence. Both of these conditional probabilities were seen to be systematically related to vectors of attributes characterizing the individuals/elements in the relevant universes, and, therefore, were seen to be representative of nonrandom events.

The reason that conditional probabilities were analyzed above owes to the major motivating factor for this study: drawing inferences about larger groups from parameters estimated on subsamples. One such group, of course, is the total population, while another group of interest is the total SMSA population. Results not presented here, but available on request, demonstrate that analyzing the probability of being included in the final sample, given that a smoking survey was administered, is significantly influenced by individual socioeconomic attributes, as expected. Thus, it seems that in extrapolating results either to the total population or to an SMSA population one must proceed with caution.

To illustrate the differences between individuals in the overall HIS, those receiving the smoking supplement, and those in the final RFF sample, we have calculated population summary statistics for the three groups. Table 2-9 presents the comparisons for adults, while Table 2-10 refers to children. One important point is worth noting. Even though our sample selection procedure is biased towards SMSA residents (no one outside an SMSA is included in our sample), and even though our procedure has a (less serious) bias toward central city residents (because of the availability of pollution monitoring data there), the characteristics of our final samples of adults and children are very similar to those of the overall population. Only with respect to SMSA and Central-city residence are there significant differences in population characteristics. Our sample does have slightly more adults and children from the northeast region of the U.S. than the HIS and slightly fewer from the south, but the age, income, education, health, labor market, and other characteristics are quite similar between the two samples.

The distinctions between the two samples arise again in Chapter 5 when we discuss the possible extrapolations of our results to larger populations.

Table 2-9. Population Summary Statistics: Adults

N	All Adults (HIS) 79,743	Smoking Survey 26,271	Final RFF Sample 14,441
	Mean (Std. Dev.)		
Age	42.54 (18.22)	42.42 (18.19)	42.98 (18.34)
Restricted activity days (two weeks)	0.847 (2.88)	0.851 (2.88)	0.882 (2.93)
Bed disability days (two weeks)	0.286 (1.55)	0.292 (1.57)	0.313 (1.63)
Work or school loss days (two weeks)	0.160 (1.09)	0.162 (1.10)	0.169 (1.12)
	Frequencies (%)		
Region of residence			
Northeast	22.64	22.61	28.40
Northcentral	26.54	26.37	26.52
South	32.18	32.34	25.46
West	18.64	18.68	19.63
Location of residence			
SMSA-Central City	28.21	28.30	44.37
SMSA-not Central city	40.54	40.43	55.63
Not in SMSA	31.26	31.27	0.00
Location of residence			
SMSA	68.75	68.73	100.00
Non SMSA non-farm	28.47	28.46	0.00
Non SMSA farm	2.79	2.81	0.00
Race			
White	88.55	88.41	85.88
Black	9.84	9.91	12.14
Other	1.60	1.68	1.98
Sex			
Male	46.75	46.71	46.53
Female	53.25	53.30	53.47
Age			
15-44	57.34	57.57	55.37
45-64	27.95	27.95	29.80
65+	14.71	14.49	14.83

Table 2-9 (continued)

	All Adults (HIS)	Smoking Survey	Final RFF Sample
N	79,743	26,271	14,441
	Mean (Std. Dev.)		
Marital status			
Married, spouse present	63.81	64.13	60.39
Widowed	7.62	7.43	7.83
Never married	20.33	20.21	22.87
Divorced	5.38	5.36	5.77
Separated	2.16	2.21	2.45
Married, spouse absent	0.69	0.65	0.70
Education (years)			
None	0.71	0.71	0.75
01-08	14.00	13.90	13.35
09-11	16.56	16.62	17.01
12	37.24	37.32	36.69
13-15	15.69	16.01	16.54
16+	14.02	13.96	14.02
Unknown	1.77	1.49	1.64
Veteran status			
Nonveteran	79.59	80.01	79.63
Peace time only	2.55	2.43	2.43
WWI	0.55	0.52	0.54
WWII	6.66	6.81	7.47
Korean War	3.21	3.07	3.17
Vietnam War	5.17	5.18	4.70
Post Vietnam	0.21	0.24	0.21
Don't know	2.05	1.76	1.85
Family income			
under $3,000	4.26	4.29	4.02
$3,000-4,999	5.90	5.83	5.75
$5,000-6,999	6.84	6.81	6.61
$7,000-9,999	8.80	8.80	8.32
$10,000-14,999	15.50	15.75	14.90
$15,000-24,999	24.77	24.73	24.40
$25,000+	24.47	24.39	25.77
Don't know	9.47	9.41	10.24
Living arrangement			
Living Alone	11.57	11.40	12.31
Living w/non relatives	2.90	2.91	3.05
Living w/spouse	63.75	64.06	60.34
Living w/relatives-other	21.78	21.63	24.30

Table 2-9 (continued)

N	All Adults (HIS) 79,743	Smoking Survey 26,271	Final RFF Sample 14,441
		Mean (Std. Dev.)	
Usual activity			
Usually working	57.23	57.63	57.08
Keeping house (female)	24.12	23.88	23.22
Retired-Health (45+ yrs)	2.46	2.61	2.41
Going to school	7.85	7.60	8.14
Something else	3.16	3.18	3.56
Retired-other (45 yrs)	4.96	4.93	5.44
Don't know	0.22	0.18	0.15
Activity limitations			
Cannot perform usual activity	4.94	5.15	5.24
Can perform usual activity, but limited in amount and kind	9.24	8.99	8.55
Can perform usual activity, but limited in outside activities	4.39	4.18	4.11
Not limited or not applicable	81.43	81.69	82.10
Health status			
Excellent	43.99	44.13	43.39
Good	39.54	39.47	40.46
Fair	11.97	11.98	11.83
Poor	3.82	3.77	3.63
Unknown	0.68	0.66	0.69

Table 2-10. Population Summary Statistics: Children

N	All Children (HIS) 30,787	Final RFF Sample 15,711
	Mean (Std. dev.)	
Age	8.49 (4.93)	8.61 (4.94)
Restricted activity days (two weeks)	0.417 (1.62)	0.416 (1.61)
Bed disability days (two weeks)	0.186 (.920)	0.191 (.944)
	Frequencies (%)	
Region of residence		
Northeast	20.50	25.83
Northcentral	27.33	28.71
South	33.88	26.81
West	18.29	18.65
Location of residence		
SMSA-Central City	26.14	43.88
SMSA-not Central City	40.55	56.62
non-SMSA	33.32	0.00
Location of residence		
SMSA	66.68	100.00
Non SMSA non-farm	30.59	0.00
Non SMSA farm	2.73	0.00
Race		
White	83.63	79.19
Black	14.43	18.59
Other	1.94	2.22
Sex		
Male	50.99	51.05
Female	49.01	48.95
Age		
Under 6	31.49	30.66
06-16	68.51	69.34
Family income		
under $3,000	3.26	3.32
$3,000-4,999	4.86	5.75
$5,000-6,999	5.83	6.15
$7,000-9,999	8.09	7.89

Table 2-10 (continued)

N	All Children (HIS) 30,787	Final RFF Sample 15,711
	Mean (Std. dev.)	
$10,000-14,999	16.42	14.92
$15,000-24,999	29.01	28.31
$25,000+	23.52	23.82
Don't know	9.02	9.84
Activity limitations		
Cannot perform usual activity	0.19	0.20
Can perform usual activity, but limited amount and kind	1.95	2.20
Can perform usual activity, but limited in outside activities	1.85	1.81
Not limited or not applicable	96.02	95.78
Health status		
Excellent	60.13	58.79
Good	34.33	33.51
Fair	4.24	4.37
Poor	0.52	0.54
Unknown	0.78	0.80

Footnotes

[1] This algorithm was developed by Professor Raymond Palmquist of the Department of Economics and Business, North Carolina State University.

[2] National Center for Health Statistics [1981], pp. 42-43.

[3] National Center for Health Statistics [1981]; National Center for Health Statistics [1975]; Bureau of the Census [1979]; National Center for Health Statistics [1979].

[4] See, for example, Lilienfeld and Lilienfeld [1980], chapters 6, 9.

[5] See, for instance, Higgins [1983].

[6] See Manning, Newhouse, and Ware [1982].

[7] See NCHS [1981], pp. 57-58.

[8] Surgeon General of the United States [1982].

[9] Warner [1978], p. 314.

[10] U.S. Environmental Protection Agency [1980], p. vii.

[11] See U.S. Environmental Protection Agency [1978], Johnson [1982].

[12] Bernard D. Goldstein, M.D., personal communication, 7/6/82.

[13] National Academy of Sciences [1981].

[14] This can be found at Tape location nos. 529-30 in the 1974 HIS.

[15] For information about the PSI, see Federal Inter-agency Task Force on Air Quality Indicators [1976].

[16] For details about this file, see U.S. Department of Commerce [1979].

[17] See Heckman [1979], pp. 153-154.

References to Chapter 2

Bureau of the Census (acting as Collecting Agent for U.S. Public Health Service), *Health Interview Survey Interviewer's Manual*, U.S. Department of Commerce. Publication no. HIS-100 (1979), 1979.

Federal Interagency Task Force on Air Quality Indicators, *A Recommended Air Pollution Index* (Washington, D.C.: U.S. Government Printing Office), 1976.

Heckman, James, "Sample Selection Bias as a Specification Error," *Econometrica*, vol. 47 (1979), pp. 153-161.

Higgins, Ian, "What Is an Adverse Health Effect?" *Journal of the Air Pollution Control Association*, vol. 33 (1983), pp. 661-663.

Johnson, Ted, "A Jump Test for Validating Hourly Average Air Quality Data," paper presented at the Annual Meeting of the Air Pollution Control Association, New Orleans, June 20-25, 1982.

Lilienfeld, A.M. and D.E. Lilienfeld, *Foundations of Epidemiology* (New York: Oxford University Press), 1980.

Manning, Willard, Joseph Newhouse, and John Ware, "The Status of Health in Demand Estimation; or, Beyond Excellent, Good, Fair, Poor," in Victor Fuchs (ed.) *Economic Aspects of Health* (Chicago: University of Chicago Press), 1982.

National Academy of Sciences, *Indoor Pollutants* (Washington, D.C.: National Academy Press), 1981.

National Center for Health Statistics, *Health Interview Survey Procedure, 1957-1974*, Department of Health and Human Services Publication no. (HRA) 75-1311, 1975.

National Center for Health Statistics, *Current Estimates from the National Health Interview Survey: United States, 1979*, Department of Health and Human Services Publication no. (PHS) 81-1564, 1981.

Surgeon General of the United States, *The Health Consequences of Smoking, Cancer*, Department of Health and Human Services Publication no. (PHS) 82-50179, 1982.

U.S. Department of Commerce (National Technical Information Service), "Bureau of Health Manpower Area Resource File User Documentation," Publication no. HRP-0901718, July 1979.

U.S. Environmental Protection Agency, *Screening Procedures for Ambient Air Quality Data*, EPA-450/2-78-037 (Research Triangle Park, N.C.), 1978.

U.S. Environmental Protection Agency, *Air Quality Data--1979 Annual Statistics*, USEPA, Office of Air Quality Planning and Standards, no. EPA-450/4-80-014, 1980.

Warner, Kenneth, "Possible Increases in The Underreporting of Cigarette Consumption," *Journal of the American Statistical Association*, vol. 73 (1978), pp. 314-318.

Chapter 3

METHODOLOGY

In the first part of this chapter, we elaborate on the brief description in Chapter 1 of our approach to health effects estimation. First we discuss measures of acute and chronic health status used in this study. It is the variations in these measures which we later attempt to explain using concentrations of ozone and other air pollutants, socioeconomic measures, and additional possible determinants. These measures of health status are the dependent variables in the series of statistical estimates presented below.

Next we identify and discuss the explanatory variables used to explain variations in each of these measures of health status. Variables used in one model might not be used in another; typically, however, a set of "core" variables turns up in each of the estimated models because they are thought to be important determinants of a number of measures of health status. Included in this discussion of independent variables are some *a priori* speculations about their likely signs in the regressions below.

We also discuss the samples over which our models are estimated. We often estimate a number of different specifications of the same basic model on one particular set of observations. It has been observed that this can cause problems when only a subset of the final results are reported.[1] We go to great lengths in this study, however, to report a wide range of results, or to make them available to those interested. Coupled with the impossibility of having used "fresh" observations for every alternative specification, we consider this to be a satisfactory justification for the approach we have employed.

Finally, we discuss briefly the estimating techniques used in this report to identify possible relationships between the dependent and independent variables. The choice of a particular technique is seen to depend in large part on the nature of the health measure being examined. In some cases fairly simple statistical techniques such as ordinary least squares (OLS) are used. In other cases more complex approaches have been adopted.

3.1 Measures of Health Status: Choice of the Dependent Variables

We have declined in this study to measure health status using the "excellent, good, fair, or poor" (EGFP) descriptors. First, any one individual might describe health status in one way at one point in time and another way at some later point even though a more "objective" measure of health status might indicate identical states. More importantly, wherever possible in this study we are interested in establishing quantitative relationships between ozone and measures of health status that might lend themselves to valuation. The distinctions possible using the "EGFP" scale do not easily lend themselves to such valuation.

Well suited to this purpose are other measures of acute health status reported in the HIS. It will be recalled that each respondent was asked the number of "restricted activity days" (RADs), "work loss (or school loss) days" (WLDs or SLDs), and "bed disability days" (BDDs) due to illness during the two weeks immediately preceding the interview date. The number of each can vary from zero to a maximum of fourteen during the two-week reference period. Because these measures are quantitative and vary continuously, and because they are amenable to economic valuation, they qualify as excellent dependent variables. We make much use of these measures in the analysis below.

As mentioned earlier, each time an acute health episode occurs--be it a RAD, a WLD, or a BDD--the cause or condition giving rise to it is ascertained and coded according to the International Classification of Diseases. Thus, during a particular individual's reference period, he or she may have one BDD due to an acute respiratory condition (asthma, bronchitis, etc.), one BDD due to an upset stomach or some other gastrointestinal problem, and one BDD because of a sprained ankle or other accidental injury. In other words, RADs, WLDs, and BDDs are broken down by specific types of injuries or diseases.

Even though respiratory health effects are those most likely to be associated with air pollution, we begin our analyses below using the number of BDDs, WLDs (or SLDs) and RADs during the reference period due to all causes (including accidents). We do this for several reasons. First, as a practical matter, respiratory disease typically accounts for a substantial fraction of the total number of BDDs, WLDs, SLDs, and RADs from all causes each year in the United States. In the 1979 HIS, for example, the percentages are 47, 37, 57, and 40, respectively.[2] Second, some evidence suggests that ozone and other air pollutants may be related to eye irritation, cardiovascular disease, susceptibility to infectious disease, and other ailments as well as respiratory disease.[3] For this reason, it is important to look beyond respiratory disease alone.

Following the analysis of RADs, WLDs, SLDs and BDDs due to all causes, we turn our attention specifically to respiratory-related impairment. The overall approach is the same: try to relate to the explanatory variables the number of days of each type of acute health impairment due to respiratory disease during the fourteen day reference period. Results are presented for separate analyses of acute respiratory disease in adults and children, and a considerable amount of space is devoted to sensitivity analyses of some of the more important findings.

Because of our interest in chronic as well as acute morbidity, we also investigate the associations between certain chronic illnesses and ozone, other air pollutants, and additional possible determinants. It is worth mentioning the measurement of chronic health status as a dependent variable in the analysis below.

Unlike the measures of acute morbidity included in the HIS, information about chronic morbidity does not lend itself so simply to continuous measurement. One is interested not so much in "how many days" of cancer or heart disease an individual has during a specified period of time, but rather whether or not the individual has such a chronic ailment. Those surveyed in the HIS are asked whether they have any limitations on their work, recreation, or other activities caused by health impairments. They are also asked the nature of any ailments reported. Thus, the HIS does identify individuals with asthma, bronchitis, emphysema, or other respiratory diseases, as well as with chronic diseases of the cardiovascular, nervous, digestive, musculoskeletal, or genito-urinary systems.

Because the data on chronic disease are of the "presence or absence" type, they are most appropriately analyzed using techniques designed for qualitative or limited dependent variables. These techniques, including probit and logistic analysis, are discussed below. With respect to the types of chronic morbidity to be considered, we concentrate primarily on possible links between ozone, other air pollutants, and chronic respiratory disease. If air pollution does impair health, this is where the impairment is most likely to occur. However, since some air pollutants have been linked to cardiovascular disease[4], we briefly explore it to see whether ozone is associated with its incidence. One problem we encounter there is the absence of data on several potentially important confounding variables. Finally, we briefly explore the relationship between annual average pollution levels and several other forms of chronic deseases.

The Appendix to Chapter 3 describes in great detail the definitions of the measures of health status used in our study. It also presents the method used to create these variables from the HIS Public Use Data Tape.

Tables 3-1 and 3-2 below summarize the frequencies of our alternative measures of acute and chronic morbidity. In Table 3-1 we compare the frequencies of morbidity in the final RFF sample of adults with the corresponding frequencies in the whole sample of adults receiving the smoking supplement.

3.2 Explanatory Variables

Here we identify and discuss the socioeconomic, pollution, meteorological and other variables used to explain variation in the measures of acute and chronic illness. Where possible we speculate about the expected signs of the independent variables in the regressions below.

Because of the role they may play in explaining acute and chronic morbidity, a number of individual socioeconomic characteristics from the HIS are included in almost every regression below. These variables--and, in parentheses, the way the variable appears in the equations that follow--include:

Table 3-1. Summary of Adults' Health Status Measures

	SMOKING SURVEY		FINAL RFF SAMPLE	
Measure	Mean	N	Mean	N
TOTRAD2W	.8508	26,271	.8818	14,441
BEDDIS2W	.2921	26,271	.3129	14,441
WKSLLD2W	.1866	16,165	.1952	8,766
RADMINOR	.3931	24,161	.3969	13,234
TRADRSP	.1741	26,271	.1674	14,441
TBEDRSP	.0775	26,271	.0137	14,441
TWSLRSP	.0500	16,165	.0557	8.766
RADRSPPS	.0426	26,271	.0414	14,441
BEDRSPPS	.0257	26,271	.0251	14,441
WSLRSPPS	.0227	16,165	.0240	8,766
PRSCHRSP	.0750	26,271	.0692	14,441
PRSCHCRD	.1049	26,271	.1071	14,441

Frequencies

Measure	Quantity	Frequency	Percent	Quantity	Frequency	Percent
TOTRAD2W	0	22657	86.243	0	12415	85.971
	1	734	2.794	1	384	2.659
	2	639	2.432	2	366	2.534
	3	403	1.534	3	220	1.523
	4	228	0.868	4	139	0.963
	5	172	0.655	5	102	0.706
	6	77	0.293	6	35	0.242
	7	201	0.765	7	110	0.762
	8	36	0.137	8	23	0.159
	9	18	0.069	9	11	0.076
	10	87	0.331	10	58	0.402
	11	14	0.053	11	5	0.035
	12	20	0.076	12	11	0.076
	13	15	0.057	13	6	0.042
	14	970	3.692	14	556	3.850

Table 3-1 (Continued).

	SMOKING SURVEY			FINAL RFF SAMPLE		
Measure	Quantity	Frequency	Percent	Quantity	Frequency	Percent
BEDDIS2W	0	24559	93.483	0	13465	93.241
	1	503	1.915	1	254	1.759
	2	359	1.367	2	224	1.551
	3	201	0.765	3	111	0.769
	4	112	0.426	4	63	0.436
	5	98	0.373	5	62	0.429
	6	31	0.118	6	23	0.159
	7	76	0.289	7	41	0.284
	8	30	0.114	8	18	0.125
	9	12	0.046	9	7	0.048
	10	36	0.137	10	23	0.159
	11	13	0.049	11	8	0.055
	12	11	0.042	12	7	0.048
	13	9	0.034	13	3	0.021
	14	221	0.841	14	132	0.914
WKSLLD2W	n/a	10106	n/a	n/a	5675	n/a
	0	15237	94.259	0	8255	94.171
	1	395	2.444	1	214	2.441
	2	192	1.188	2	103	1.175
	3	84	0.520	3	42	0.479
	4	37	0.229	4	21	0.240
	5	61	0.377	5	38	0.433
	6	16	0.099	6	10	0.114
	7	17	0.105	7	10	0.114
	8	9	0.056	8	4	0.046
	9	5	0.031	9	3	0.034
	10	85	0.526	10	53	0.605
	11	2	0.012	12	5	0.057
	12	10	0.062	14	8	0.091
	13	1	0.006			
	14	14	0.087			
RADMINOR	n/a	2110	n/a	n/a	1207	n/a
	0	22657	93.775	0	12415	93.811
	1	277	1.146	1	144	1.088
	2	294	1.217	2	161	1.217
	3	182	0.753	3	95	0.718
	4	92	0.381	4	54	0.408
	5	50	0.207	5	30	0.227
	6	25	0.103	6	10	0.076
	7	91	0.377	7	47	0.355
	8	6	0.025	8	3	0.023
	9	4	0.017	9	3	0.023
	10	37	0.153	10	26	0.196
	11	1	0.004	12	3	0.023
	12	6	0.025	13	2	0.015
	13	2	0.008	14	241	1.821

Table 3-1 (Continued).

	SMOKING SURVEY			FINAL RFF SAMPLE		
Measure	**Quantity**	**Frequency**	**Percent**	**Quantity**	**Frequency**	**Percent**
TRADRSP	0	25153	95.744	0	13843	95.859
	1	328	1.249	1	165	1.143
	2	262	0.997	2	141	0.976
	3	141	0.537	3	87	0.602
	4	85	0.324	4	48	0.332
	5	61	0.232	5	33	0.229
	6	18	0.069	6	9	0.062
	7	62	0.236	7	32	0.222
	8	9	0.034	8	4	0.028
	9	2	0.008	9	2	0.014
	10	16	0.061	10	11	0.076
	11	1	0.004	11	1	0.007
	12	4	0.015	12	3	0.021
	13	2	0.008			
	14	127	0.483	14	62	0.429
TBEDRSP	0	25597	97.434	0	14078	97.486
	1	265	1.009	1	128	0.886
	2	174	0.662	2	102	0.706
	3	85	0.324	3	52	0.360
	4	34	0.129	4	20	0.138
	5	34	0.129	5	22	0.152
	6	6	0.023	6	5	0.035
	7	22	0.084	7	12	0.083
	8	8	0.030	8	4	0.028
	9	3	0.011	9	2	0.014
	10	4	0.015			
	12	2	0.008	12	1	0.007
	13	1	0.004			
	14	36	0.137	14	15	0.104
TWSLRSP	n/a	10106	n/a	n/a	5675	n/a
	0	15798	97.730	0	8556	97.604
	1	192	1.188	1	104	1.186
	2	84	0.520	2	48	0.548
	3	36	0.223	3	22	0.251
	4	16	0.099	4	10	0.114
	5	19	0.118	5	12	0.137
	6	3	0.019	6	2	0.023
	7	4	0.025	7	4	0.046
	8	1	0.006	8	1	0.011
	10	10	0.062	10	6	0.068
	14	2	0.012	14	1	0.011

Table 3-2. Summary of Childrens' Health Status Measures

FINAL RFF SAMPLE

Measure	Mean	N
TOTRAD2W	.4156	15,711
BEDDIS2W	.1913	15,711
WKSLLD2W	.2615	7,972
RADMINOR	.1159	14,283
TRADRSP	.2191	15,711
TBEDRSP	.1200	15,711
TWSLRSP	.1618	7,972
RADRSPPS	.0707	15,711
BEDRSPPS	.0465	15,711
WSLPSPPS	.0664	7,972

Measure	Quantity	Frequency	Percent
TOTRAD2W	0	13808	87.887
	1	585	3.724
	2	460	2.928
	3	305	1.941
	4	147	0.936
	5	112	0.713
	6	36	0.229
	7	90	0.573
	8	18	0.115
	9	10	0.064
	10	14	0.089
	11	4	0.025
	12	4	0.025
	13	2	0.013
	14	116	0.738

Table 3-2 (Continued).

FINAL RFF SAMPLE

Measure	Quantity	Frequency	Percent
BEDDIS2W	0	14578	92.788
	1	429	2.731
	2	310	1.973
	3	171	1.088
	4	74	0.471
	5	43	0.274
	6	11	0.070
	7	50	0.318
	8	10	0.064
	9	6	0.038
	10	5	0.032
	11	1	0.006
	14	23	0.146
WKSLLD2W	n/a	7739	n/a
	0	7130	89.438
	1	362	4.541
	2	225	2.822
	3	97	1.217
	4	46	0.577
	5	64	0.803
	6	8	0.100
	7	6	0.075
	8	2	0.025
	9	3	0.038
	10	14	0.176
	12	2	0.025
	13	1	0.013
	14	12	0.151
RADMINOR	n/a	1428	n/a
	0	13808	96.674
	1	142	0.994
	2	120	0.840
	3	83	0.581
	4	36	0.252
	5	24	0.168
	6	5	0.035
	7	18	0.126
	8	5	0.035
	9	2	0.014
	10	2	0.014
	11	2	0.014
	14	36	0.252

Table 3-2 (Continued).

FINAL RFF SAMPLE

Measure	Quantity	Frequency	Percent
TRADRSP	0	14601	92.935
	1	351	2.234
	2	293	1.865
	3	190	1.209
	4	80	0.509
	5	62	0.395
	6	18	0.115
	7	50	0.318
	8	5	0.032
	9	3	0.019
	10	6	0.018
	11	1	0.006
	12	1	0.006
	13	2	0.013
	14	48	0.306
TBEDRSP	0	14980	95.347
	1	276	1.757
	2	206	1.311
	3	117	0.745
	4	44	0.280
	5	27	0.172
	6	8	0.051
	7	27	0.172
	8	6	0.018
	9	2	0.013
	10	4	0.015
	11	1	0.006
	14	13	0.083
TWSLRSP	n/a	7739	n/a
	0	7443	93.364
	1	217	2.722
	2	158	1.982
	3	62	0.778
	4	24	0.301
	5	41	0.514
	6	6	0.075
	7	2	0.025
	8	2	0.025
	9	1	0.013
	10	7	0.088
	13	1	0.013
	14	8	0.100

o Age (AGE)--Age may play an important role in explaining both acute and chronic morbidity. When it is entered linearly, one might hypothesize that the older the individual the greater the number of days of acute illness (or the greater the likelihood of having some acute morbidity) he or she will have during a two-week period. Similarly, older individuals are hypothesized to be more likely to have a chronic respiratory or other illness subject to aggravation by air pollution. In both cases, the rationale is the same: over time, an individual's stock of "health capital" wears down, making him or her more susceptible to acute or chronic morbidity.

However, one need not confine the investigation to linear functional forms. In some cases, one might expect a quadratic (or U-shaped) relationship between age and morbidity. To explore for this possible non-linear relationship between age and health, we use both age and its square of age (AGESQ) in several regressions.

o Income (INCOMCON)--Another potentially important influence on acute and chronic morbidity is income. Wealthier individuals are able to purchase higher quality health care as well as other goods and services likely to contribute to good health, including better foods, more leisure time, and so on. Thus, we would expect a negative association between income and morbidity.

The only information sought by HIS from the respondents is an indication of which of eleven intervals their annual household income falls into (see Appendix A at the end of this report for details). Rather than work with ten dummy variables in each regression, we converted the interval data into a continuous measure of income using the midpoints of the income ranges and assigning an income of $30,000 to everyone in the "$25,000 and greater" category. It is this "continuous" measure of income that is used in the regressions.

o Education (EDCOMCON)--Education may affect health a number of ways: by familiarizing individuals with the early signs and symptoms of disease, by familiarizing them with preventive measures, and by impressing upon them the importance of periodic medical checkups. Also, education might also help people recover more quickly and easily from diseases they do contract--make them more efficient "producers of health," in other words. Therefore, years of education completed ought to be negatively associated with acute and chronic morbidity. (This was also converted to a continuous measure using interval data from the HIS. Again, see the questionnaire in the Appendix to Chapter 2.)

o Race (RACEW1B0)--Although it is not clear which direction such a relation would take, race could play a significant role in explaining variations in health status. RACEW1B0 is a dummy variable taking on the value unity if the respondent is white, zero if black.

o Sex (SEXM1F0)--Differences between the sexes might also account for variations in acute and chronic morbidity. To account for this possibility, the dummy variable SEXM1F0 is included which takes on the value unity if the individual is a male, zero if female.

o Marital status (MARY1N0)--The presence of a spouse may enable one to recover more quickly or easily from an illness, and some evidence suggests that married individuals may be healthier for psychological reasons as well. To test for these effects the dummy variable MARY1N0 takes on the value unity if the respondent is married, zero otherwise.

- General condition (FAT)--As discussed earlier, it would be desirable to have information about individuals' exercise habits and physical condition. The HIS provides no such information but does include the height and weight of all adults. As a crude proxy for physical condition, we have created the variable FAT, equal to weight in pounds divided by height in inches. Since overly thin individuals may, along with the obese, be prone to illness, we generally use a quadratic form including both FAT and its square, FATSQ.

- Smoking (SMOKY1N0, NCIGSDYN)--No environmental influence has been more clearly linked to acute and chronic morbidity than personal smoking. There is even evidence that proximity to smokers may be injurious to health. Accordingly we measure smoking in one of two ways in the analyses below. SMOKY1N0 is a dummy variable taking on the value unity if the respondent currently smokes occasionally or regularly, zero if not. We also measure smoking on the basis of the number of cigarettes an individual usually smokes each day, NCIGSDYN. *A priori* we expect a positive association between smoking and acute and chronic illness.

- Chronic illness (CHRLMDUM)--When analyzing the determinants of acute morbidity (work days lost, bed disability days, etc.), it is important to know whether an individual also has any chronic illnesses. If so, we would expect him or her to have more days of acute morbidity during a specified period than if such a condition(s) was missing. To account for this possibility, we include in our analysis of acute morbidity the dummy variable CHRLMDUM which takes on the value unity if the individual reports having a chronic illness, zero if otherwise. In our analysis of chronic morbidity, a variable like this one becomes the dependent variable and hence is not included among the regressors.

Several other individual or household characteristics are available from the HIS. Although of less potential interest than those already considered, they do appear in some of the reported regressions. They include:

o Veteran status (VETY1N0)--A dummy variable taking the value unity if the respondent is a veteran of the armed services, zero if not.

o Central city residence (CITYI100)--To see whether residents of central cities are more or less healthy than others, we include in some regressions a dummy variable taking the value unity if the respondent lives in a central city, zero if not.

o Paid sick leave (INDOCCAV)--As discussed previously, the amount of paid sick leave to which an individual is entitled may influence the amount of reported work loss due to illness. INDOCCAV measures the average number of days of annual paid sick leave in 1974 received by someone who worked in the same profession and industry as the respondent. This data comes from a supplement to the 1974 HIS (see Chapter 2 for the precise reference).

o Profession (BLUE)--Individuals working in certain occupations may be exposed to more accident and health risks than those in other occupations, and thus experience more acute and/or chronic illness. To test this hypothesis, we include in some regressions a dummy variable BLUE taking the value unity if the respondent is employed as a craftsman or kindred worker, equipment operator, farm or other laborer, or service worker, and zero if employed as a professional technical or kindred worker, a manager or administrator, salesperson, clerical or kindred worker.

o Household density (CROWDING)--The closer the contact between individuals, the more likely it is that a communicable disease will be transmitted between them. To allow for this possibility, we created the variable CROWDING by dividing the number of persons in a household by the number of rooms in the dwelling. Our expectation is that this measure should vary directly with acute and chronic morbidity (the lower the density, the less likely the incidence of morbidity).

o Ozone and other air pollutants (see Chapter 2 and its appendix)--The measures of air pollution used in our analysis of acute and chronic morbidity are defined and described in considerable detail in Chapter 2 and its Appendix. Because of the large number of different pollution measures we employ, they are not discussed here. Readers are encouraged to familiarize themselves with the pollution measures before proceding with the analyses. All air pollution variables may be expected to vary directly with acute and chronic morbidity.

Measures of meteorological conditions appear in virtually every regression. Among the most important of these are the following:

- Temperature (AVMAXTMP)--Temperature, particularly extreme temperatures, can influence acute morbidity. Accordingly, several different measures of temperature are employed in our analyses. AVMAXTMP is the average of the highest daily temperatures in degrees Fahrenheit during the two-week reference period for which data are available on acute morbidity.

The relationship between temperature and acute morbidity is a complex one. During the summer months, for example, one might expect hotter temperatures to be associated with more acute illness. During the winter months, however, warmer temperatures might signify more temperate climates where illness might be expected to be lower. For this reason, we created two interactive/dummy variables, DMAXTEMP and DMINTEMP. The former measures maximum temperature for individuals interviewed in the second and third quarters of 1979 (April-October). The latter is the operative temperature measure during quarters 1 and 4 and measures the average minimum daily temperature. This approach allows one to test whether high temperatures are the relevant measure in summer months, while low temperatures influence acute morbidity in winter months. Appropriate annual measures are used in our analysis of chronic morbidity.

o Precipitation (AVPRECIP)--Snow or rainfall might also increase the likelihood or amount of acute and chronic morbidity, or influence work or school attendance even when an illness is present. To test this possibility for acute illness, we include a measure of the average daily amount of precipitation in inches during the two-week reference period. As in the case of temperature, annual measures of precipitation are included in the analysis of chronic morbidity.

o Humidity (HUMIDRF)--This variable measures the relative humidity during the two-week reference period. As in the case of temperature,, annual measures of precipitation are included in the analysis of chronic morbidity.

Several other kinds of variables receiving mention in Chapter 2 are used in our analysis. They include:

o Indoor air pollution (PGASPIPD, PGASALL)--Indoor exposures to certain air pollutants can be high, and gas stoves are reported to be responsible for high indoor levels of several of these pollutants, in particular nitrogen dioxide. Therefore, PGASPIPD and PGASALL--respectively, the percentage of all occupied housing units in the individual's SMSA that use piped natural gas as a cooking fuel, and the percentage cooking with piped as well as bottled natural gas--are sometimes included as independent variables in our analysis of chronic morbidity. A priori, the higher this percentage, the more likely the presence of a chronic respiratory disease.

- Air quality information (PSIY1N0)--A dummy variable taking on the value unity if the PSI (Pollutant Standard Index) is reported regularly in an individual's SMSA, zero if not. As discussed above, the expected sign of this variable is ambiguous.

- Availability of medical care (PCDOCN)--This variable measures the per capita number of doctors in an individual's SMSA, and is intended to serve as a rough measure of the availability of medical care.

- Airborne allergens (POLLENRF, POLLENAN)--Both acute and chronic morbidity may be influenced by natural substances in ambient air to which people are allergic. To test this hypothesis, we include the variables POLLENRF and POLLENAN which measure, respectively, the ragweed concentration during the two-week reference period and the annual average pollen concentration.

The variables discussed above are listed in Table 3-3. Beside each variable we indicate its expected sign in regressions explaining acute and/or chronic morbidity.

Tables 3-4 and 3-5 below present the mean and extreme values for these independent variables for both the adults' and children's data sets. They indicate substantial variation in the independent variables. For instance, the mean value for O3NR01, the average daily maximum one-hour ozone concentration during the two-week reference period, was 0.045 parts-per-million--well below the current ozone

Table 3-3. Independent Variables	Expected Sign
Age (AGE)	+
Income (INCOMCON)	-
Education (EDCOMCON)	-
Race (RACEW1B0)	?
Sex (SEXM1F0)	?
Marital status (MARY1N0)	-
Physical condition (FAT)	+
Smoking (SMOKY1N0, NCIGSDYN)	+
Chronic illness (CHRLMDUM)	+
Veteran status (VETY1N0)	?
Central city dweller (CITYI100)	?
Paid sick leave (INDOCCAV)	?
Profession (BLUE)	+(?)
Household density (CROWDING)	+
Air pollutants (see chapter 2)	+
Temperature (AVMAXTMP)	+(?)
Precipitation (AVPRECIP)	+
Humidity (HUMIDRF)	+(?)
Indoor air pollution (PGASPIPD, PGASALL)	+
Air quality information (PSIY1N0)	?
Availability of medical care (PCDOCN)	-
Airborne allergens (POLLENRF, POLLENAN)	+

Table 3-4. Mean and Extreme Values for Independent Variables: Adults Data Set

VARIABLE	N	MEAN	STD DEV	MINIMUM	MAXIMUM
O3NR01	11241	0.04497087	0.02768169	0	0.25070000
O3NR03	11241	0.07834773	0.05015616	0	0.43000000
O3NR04	11241	0.02070056	0.01276845	0	0.11660000
SPNR01	13772	70.12847836	33.75064942	9.99600000	284.00400000
SPNR03	13772	87.95955417	45.24892307	9.99600000	417.08800000
SPNR04	13772	70.12042318	33.74111915	9.99600000	284.00400000
S4NR01	7524	10.58428655	5.55484933	0	52.13600000
S4NR03	7524	12.54928814	6.95435699	0	54.88000000
S4NR04	7524	10.58428655	5.55484933	0	52.13600000
S2NR01	10788	66.20793882	52.40448212	0	760.08800000
S2NR03	10788	165.56160497	148.35212877	0	2346.90400000
S2NR04	10788	25.72125732	23.33196969	0	274.20400000
N2NR01	8555	112.38310088	59.84677997	0	435.31600000
N2NR03	8555	194.13148194	115.87018889	0	921.20000000
N2NR04	8555	64.78632800	36.08973469	0	268.71600000
CONR01	10962	3.71809747	3.00099905	0	26.58330000
CONR03	10962	7.08464149	5.60347289	0	45.00000000
CONR04	10962	1.67090242	1.40141104	0	11.40360000
AVMAXTMP	14155	65.10433466	18.33135724	4.21428571	106.35714286
AVMINTMP	14155	47.15640107	17.41837652	-13.07142857	81.57142857
DMAXTEMP	14155	33.29053843	39.78042356	0	106.35714286
DMINTEMP	14155	21.59523137	21.52006128	-13.07142857	73.28571429
HUMIDPF	14089	60.04905998	9.11759623	11.00000000	89.00000000
AVPRECIP	14155	0.10952551	0.10934349	0	1.04071429
EDCONCON	14204	11.80164038	3.23589527	0	18.00000000
INCONCON	12963	17695.98086863	9514.78016312	500.00000000	30000.00000000
RACE=1NO	14155	0.87615684	0.32941419	0	1.00000000

Reproduced from best available copy.

Table 3-4. (continued)

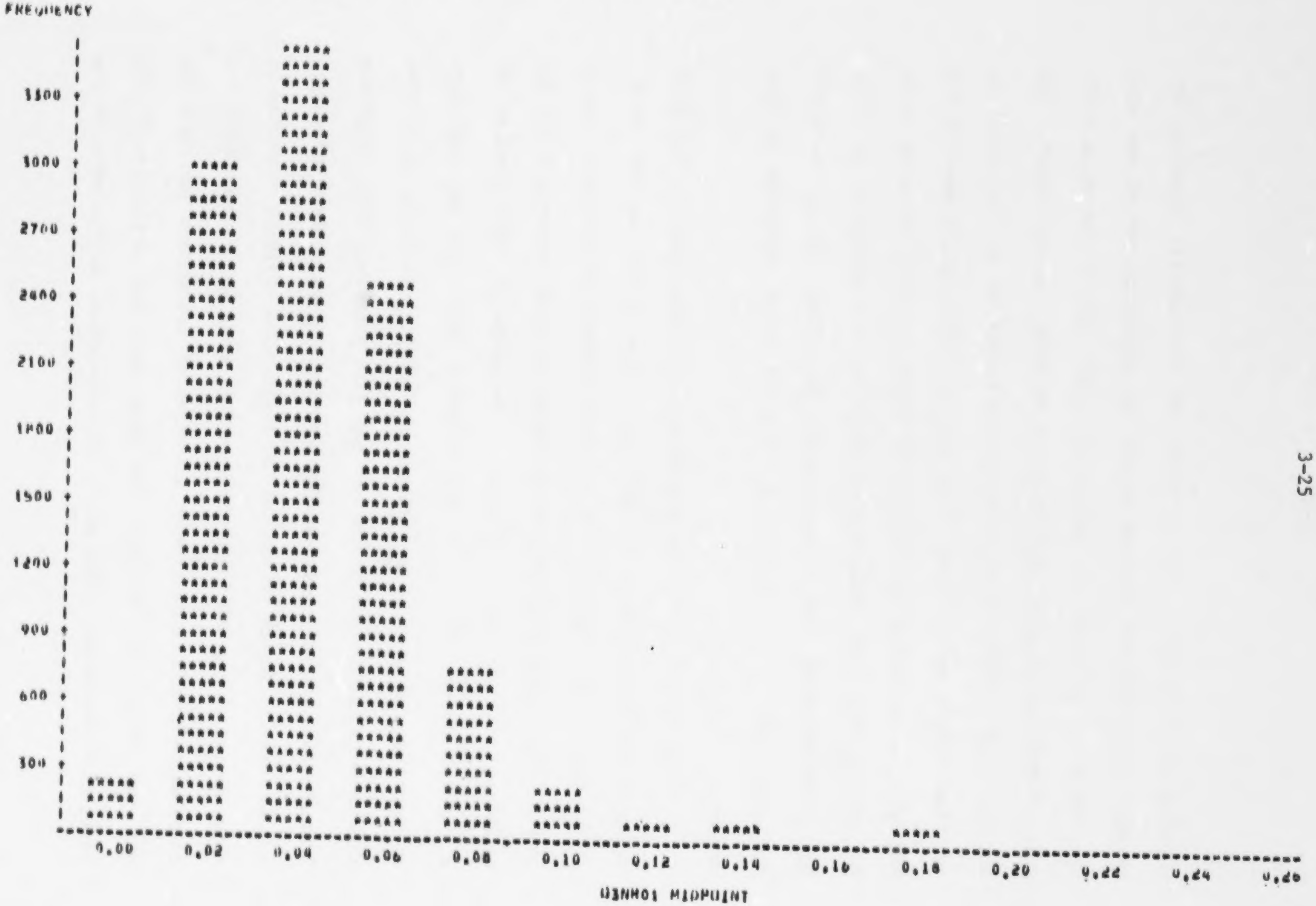
VARIABLE	N	MEAN	STD DEV	MINIMUM	MAXIMUM
SEXMIFO	14441	0.46527249	0.49880981	0	1.00000000
AGE	14441	42.97610949	18.33753012	17.00000000	99.00000000
AGESQ	14441	2183.18772938	1740.54920158	289.00000000	9801.00000000
FAT	14095	2.30001806	0.43012148	1.28985507	5.59183673
FATSQ	14095	5.47507445	2.16783306	1.66372611	31.26863807
MARYIND	14441	0.61089952	0.48756308	0	1.00000000
VETYIND	14174	0.18865528	0.39124835	0	1.00000000
SMOKYIND	13547	0.35875101	0.47965165	0	1.00000000
NCIGSDYN	13231	6.96039604	12.15097534	0	98.00000000
CHRLMOUM	14441	0.17900422	0.38336913	0	1.00000000
CITYI100	14441	0.44377658	0.49684155	0	1.00000000
PSIYIND	11851	0.84992018	0.36104048	0	1.00000000
BLUE	9066	0.46117362	0.49851773	0	1.00000000
INDOCCAV	8345	4.45099880	2.88401749	0.48000000	13.18000000
CROWDING	14171	0.59332871	0.33126706	0.11111111	8.00000000
POLLENHF	5014	1.56621478	6.29491465	0	99.70000000
PCDOCN	14441	0.00204522	0.00067915	0.00064510	0.00635100

Reproduced from best available copy.

Table 3-5. Mean and Extreme Values of Independent Variables: Children's Data Set

VARIABLE	N	MEAN	STD DEV	MINIMUM	MAXI
O3NR01	12063	0.04564713	0.02783493	0	0.25070
O3NR03	12063	0.07970041	0.05043564	0	0.43000
O3NR04	12063	0.02107658	0.01290666	0	0.10600
SPNR01	14946	71.17668005	33.42995016	12.43000000	260.09200
SPNR03	14946	89.28177680	44.99632539	13.13200000	417.00000
SPNR04	14946	71.17037227	33.42329608	12.43000000	260.09200
S4NR01	8007	10.67694517	5.79842457	0	52.13000
S4NR03	8007	12.58673960	7.12360961	0	54.60000
S4NR04	8007	10.67694517	5.79842457	0	52.13000
S2NR01	11585	66.63862961	51.64177809	0	570.30000
S2NR03	11585	180.25131118	142.20453962	0	1500.00000
S2NR04	11585	26.14420060	25.12493444	0	274.20400
N2NR01	8996	111.15356959	58.99923379	0	435.31000
N2NR03	8996	194.92186483	117.90630396	0	921.20000
N2NR04	8996	63.40795687	35.01510949	0	268.71000
CONR01	11745	3.53889072	3.02866545	0	26.50300
CONR03	11745	6.99934525	5.73542583	0	45.00000
CONR04	11745	1.63395251	1.44679191	0	12.57170
AVMAXTMP	15345	65.04237071	18.70600616	-.21428571	100.35714
AVMINTMP	15345	47.09709072	17.75751479	-13.07142857	81.57142
DMAXTEMP	15345	33.79356324	40.05236144	0	100.35714
DMINTEMP	15345	21.16335242	21.34689047	-13.07142857	73.28571
HUMIDAF	15287	60.13040609	9.26983846	11.00000000	89.00000
AVPRECIP	15345	0.11069036	0.11279767	0	1.06071
INCOMCON	14165	17750.44122838	9174.21098823	500.00000000	30000.000000
RACEX130	15342	0.80992058	0.39237042	0	1.000000
SEXMIF	15711	0.51047037	0.49990627	0	1.000000
AGE	15711	8.60715422	4.93885822	0	16.000000
AGESQ	15711	98.47387191	84.04597152	0	256.000000
CHPLMOUM	15711	0.04218973	0.20105096	0	1.000000
CITYT100	15711	0.43377252	0.49561029	0	1.000000
CROWDING	15589	0.44881459	0.38599014	0.11111111	8.000000
POLLENAF	5468	1.79757586	7.21190425	0	99.700000
PCPPC	15711	0.00200376	0.00067073	0.00004515	0.006351

Figure 3-1. Frequency Distribution of O3NR01 (Average Daily 1-Hour Maximum Ozone Concentration During Two-week Reference Period, 14,441 Adults).



standard of 0.12 ppm. For a number of individuals, however, the average daily one-hour maximum during the reference period was zero ppm, while for at least one individual O3NR01 took on the value 0.251 ppm. Figure 3-1 presents the frequency distribution for O3NR01. The highest single hourly ozone reading experienced by any individual in our final sample was 0.43 ppm (indicated by O3NR03, which measures the highest reading during the two-week reference period). Although we do not discuss them here, considerable variation also exists in the other pollution measures, the meteorological variables, the socioeconomic characteristics of the population, and the other variables we have created.

A final remark about the independent variables used in this study concerns the correlations between them. When close (but not exact) relationships exist between two or more independent variables, it is possible to obtain least-squares estimates of their parameters but the estimates will be very difficult to interpret. This problem of multicollinearity, as it is called, often arises in air pollution epidemiology because concentrations of certain pollutants often move together. (This is due in part to the fact that they often originate from the same sources, e.g., automobile exhaust, power plant emissions, and so forth.) Before discussing the statistical analysis conducted in this study, we present in Tables 3-6 and 3-7 respectively the correlation coefficients between all pairs of independent variables for all the adults and children for whom data are available on two particular variables. (The matrix of correlation coefficients for the

Table 3-6. Correlation Coefficients for Sample of Adults (N = 14,441)

```
                          CORRELATION COEFFICIENTS / PROB > |R| UNDER HO:RHO=0 / NUMBER OF OBSERVATIONS

          O3HR01    O3HR03    O3HR04    SPHR01    SPHR03    SPHR04    S4HR01    S4HR03    S4HR04    S2HR01    S2HR03    S2HR04    N2HR01

O3HR01   1.00000   0.92692   0.90897   0.29934   0.21713   0.29954   0.11875   0.11114   0.11875  -0.19697  -0.17270  -0.23424   0.28101
          0.0000    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001
           11241     11241     11241     11198     11198     11198      6575      6575      6575      9764      9764      9764      8168
O3HR03   0.92692   1.00000   0.79768   0.28597   0.21083   0.28621   0.12977   0.15359   0.12977  -0.19400  -0.17226  -0.22276   0.31661
          0.0001    0.0000    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001
           11241     11241     11241     11198     11198     11198      6575      6575      6575      9764      9764      9764      8168
O3HR04   0.90897   0.79768   1.00000   0.20021   0.12997   0.20035   0.09969   0.07033   0.09969  -0.20109  -0.14893  -0.23641   0.14768
          0.0001    0.0001    0.0000    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001
           11241     11241     11241     11198     11198     11198      6575      6575      6575      9764      9764      9764      8168
SPHR01   0.29934   0.28597   0.20021   1.00000   0.91771   0.99996   0.17001   0.15689   0.17001   0.04037   0.00290   0.00473   0.39405
          0.0001    0.0001    0.0001    0.0000    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.7631    0.6237    0.0001
           11198     11198     11198     13772     13772     13772      7510      7510      7510     10780     10780     10780      8535
```

Table 3-6. (continued)

CORRELATION COEFFICIENTS / PROB > |R| UNDER HO:RHO=0 / NUMBER OF OBSERVATIONS

	O3HR01	O3HR03	O3HR04	SPHR01	SPHR03	SPHR04	S4HR01	S4HR03	S4HR04	S2HR01	S2HR03	S2HR04	N2HR01
SPHR03	0.21713 0.0001 11198	0.21083 0.0001 11198	0.12997 0.0001 11198	0.91771 0.0001 13772	1.00000 0.0000 13772	0.91763 0.0001 13772	0.15179 0.0001 7510	0.17596 0.0001 7510	0.15179 0.0001 7510	0.06688 0.0001 10780	0.03232 0.0008 10780	0.04097 0.0001 10780	0.39740 0.0001 8535
SPHR04	0.29954 0.0001 11198	0.28621 0.0001 11198	0.20035 0.0001 11198	0.99996 0.0001 13772	0.91753 0.0001 13772	1.00000 0.0000 13772	0.16995 0.0001 7510	0.15687 0.0001 7510	0.16995 0.0001 7510	0.04003 0.0001 10780	0.00259 0.7878 10780	0.00442 0.6464 10780	0.39413 0.0001 8535
S4HR01	0.11875 0.0001 6575	0.12977 0.0001 6575	0.09969 0.0001 6575	0.17001 0.0001 7510	0.15179 0.0001 7510	0.16995 0.0001 7510	1.00000 0.0000 7524	0.90069 0.0001 7524	1.00000 0.0001 7524	0.22364 0.0001 6349	0.17903 0.0001 6349	0.25720 0.0001 6349	0.02122 0.1248 5232
S4HR03	0.11114 0.0001 6575	0.15169 0.0001 6575	0.07033 0.0001 6575	0.18689 0.0001 7510	0.17596 0.0001 7510	0.15687 0.0001 7510	0.90069 0.0001 7524	1.00000 0.0000 7524	0.90069 0.0001 7524	0.16413 0.0001 6349	0.12741 0.0001 6349	0.20682 0.0001 6349	0.15698 0.0001 5232
S4HR04	0.11875 0.0001 6575	0.12977 0.0001 6575	0.09969 0.0001 6575	0.17001 0.0001 7510	0.15179 0.0001 7510	0.16995 0.0001 7510	1.00000 0.0001 7524	0.90069 0.0001 7524	1.00000 0.0000 7524	0.22364 0.0001 6349	0.17903 0.0001 6349	0.25720 0.0001 6349	0.02122 0.1248 5232
S2HR01	-0.19897 0.0001 9764	-0.19480 0.0001 9764	-0.20109 0.0001 9764	0.04037 0.0001 10780	0.06688 0.0001 10780	0.04003 0.0001 10780	0.22364 0.0001 6349	0.16413 0.0001 6349	0.22364 0.0001 6349	1.00000 0.0000 10788	0.85378 0.0001 10788	0.84767 0.0001 10788	-0.05014 0.0001 7794
S2HR03	-0.17270 0.0001 9764	-0.17726 0.0001 9764	-0.14893 0.0001 9764	0.00290 0.7631 10780	0.03232 0.0008 10780	0.00259 0.7878 10780	0.17903 0.0001 6349	0.12741 0.0001 6349	0.17903 0.0001 6349	0.85378 0.0001 10788	1.00000 0.0000 10788	0.59287 0.0001 10788	-0.09936 0.0001 7794
S2HR04	-0.23474 0.0001 9764	-0.22276 0.0001 9764	-0.23661 0.0001 9764	0.00473 0.6237 10780	0.04097 0.0001 10780	0.00442 0.6464 10780	0.25720 0.0001 6349	0.20682 0.0001 6349	0.25720 0.0001 6349	0.84767 0.0001 10788	0.59287 0.0001 10788	1.00000 0.0000 10788	-0.02934 0.0096 7794
N2HR01	0.28101 0.0001 8168	0.31661 0.0001 8168	0.14768 0.0001 8168	0.39405 0.0001 8535	0.39740 0.0001 8535	0.39413 0.0001 8535	0.02122 0.1248 5232	0.15698 0.0001 5232	0.02122 0.1248 5232	-0.05014 0.0001 7794	-0.09936 0.0001 7794	-0.02934 0.0096 7794	1.00000 0.0000 8555
N2HR03	0.23570 0.0001 8168	0.28558 0.0001 8168	0.11077 0.0001 8168	0.35045 0.0001 8535	0.35978 0.0001 8535	0.35036 0.0001 8535	0.01966 0.1551 5232	0.16029 0.0001 5232	0.01966 0.1551 5232	-0.03390 0.0028 7794	-0.05617 0.0001 7794	-0.02850 0.0119 7794	0.90509 0.0001 8555
N2HR04	0.18416 0.0001 8168	0.23099 0.0001 8168	0.05032 0.0001 8168	0.34742 0.0001 8535	0.36103 0.0001 8535	0.34748 0.0001 8535	0.04659 0.0007 5232	0.17785 0.0001 5232	0.04659 0.0007 5232	-0.01694 0.1348 7794	-0.08465 0.0001 7794	0.03985 0.0004 7794	0.95640 0.0001 8555
CONR01	-0.00256 0.7966 10100	0.00477 0.6318 10100	-0.08982 0.0001 10100	0.21101 0.0001 10934	0.22987 0.0001 10934	0.21103 0.0001 10934	-0.07036 0.0001 6624	0.01983 0.1067 6624	-0.07036 0.0001 6624	-0.03179 0.0018 9645	-0.05894 0.0001 9645	-0.02071 0.0420 9645	0.39543 0.0001 8287
CONR03	-0.05375 0.0001 10100	-0.04083 0.0001 10100	-0.13302 0.0001 10100	0.16531 0.0001 10934	0.19727 0.0001 10934	0.16517 0.0001 10934	-0.06502 0.0001 6624	0.01594 0.1945 6624	-0.06502 0.0001 6624	-0.00659 0.5178 9645	-0.02743 0.0071 9645	-0.00751 0.4607 9645	0.29751 0.0001 8287

Table 3-6. (continued)

CORRELATION COEFFICIENTS / PROB > |R| UNDER HO:RHO=0 / NUMBER OF OBSERVATIONS

	O3HR01	O3HR03	O3HR04	SPHR01	SPHR03	SPHR04	S4HR01	S4HR03	S4HR04	S2HR01	S2HR03	S2HR04	H2HR01
COHR04	0.05949 0.0001 10100	0.07131 0.0001 10100	-0.03780 0.0001 10100	0.21081 0.0001 10934	0.22075 0.0001 10934	0.21090 0.0001 10934	-0.04153 0.0007 6624	0.05323 0.0001 6624	-0.04151 0.0007 6624	-0.04479 0.0001 9645	-0.08248 0.0001 9645	-0.01063 0.2967 9645	0.43523 0.0001 8287
AVMAXTMP	0.53317 0.0001 11241	0.52935 0.0001 11241	0.49404 0.0001 11241	0.20982 0.0001 13739	0.14624 0.0001 13739	0.21003 0.0001 13739	0.02931 0.0110 7518	0.03188 0.0058 7518	0.02931 0.0110 7518	-0.35171 0.0001 10788	-0.23479 0.0001 10788	-0.45246 0.0001 10788	0.08481 0.0001 8555
AVMINTMP	0.53371 0.0001 11241	0.54489 0.0001 11241	0.47866 0.0001 11241	0.15259 0.0001 13739	0.09166 0.0001 13739	0.15279 0.0001 13739	0.05864 0.0001 7518	0.06925 0.0001 7518	0.05864 0.0001 7518	-0.36709 0.0001 10788	-0.25816 0.0001 10788	-0.43773 0.0001 10788	0.10962 0.0001 8555
DMAXTEMP	0.51440 0.0001 11241	0.52539 0.0001 11241	0.49424 0.0001 11241	0.12991 0.0001 13739	0.06275 0.0001 13739	0.12999 0.0001 13739	0.19349 0.0001 7518	0.17368 0.0001 7518	0.19349 0.0001 7518	-0.19798 0.0001 10788	-0.12372 0.0001 10788	-0.24764 0.0001 10788	-0.03787 0.0005 8555
DMINTEMP	-0.33831 0.0001 11241	-0.35112 0.0001 11241	-0.34033 0.0001 11241	-0.04644 0.0001 13739	0.00125 0.8835 13739	-0.04640 0.0001 13739	-0.24462 0.0001 7518	-0.20745 0.0001 7518	-0.24462 0.0001 7518	-0.00420 0.6631 10788	-0.02130 0.0269 10788	0.00210 0.8271 10788	0.12348 0.0001 8555
HUMIDIT	-0.13215 0.0001 11152	-0.11550 0.0001 11152	-0.12895 0.0001 11152	-0.22564 0.0001 13673	-0.19222 0.0001 13673	-0.22577 0.0001 13673	0.13899 0.0001 7466	0.12728 0.0001 7466	0.13899 0.0001 7466	0.07197 0.0001 10713	0.05504 0.0001 10713	0.12062 0.0001 10713	-0.01544 0.1552 8490
AVPRECIP	-0.18083 0.0001 11241	-0.15875 0.0001 11241	-0.12789 0.0001 11241	-0.24960 0.0001 13739	-0.20708 0.0001 13739	-0.24976 0.0001 13739	0.09979 0.0001 7518	0.11178 0.0001 7518	0.09979 0.0001 7518	-0.06109 0.0001 10788	-0.01673 0.0822 10788	-0.05051 0.0001 10788	-0.22783 0.0001 8555
EDCONCUN	-0.00771 0.4177 11050	-0.01032 0.2779 11050	-0.00935 0.3258 11050	-0.06004 0.0001 13542	-0.05074 0.0001 13542	-0.05987 0.0001 13542	-0.02824 0.0150 7415	-0.03053 0.0086 7415	-0.02824 0.0160 7415	-0.03263 0.0008 10597	-0.02607 0.0073 10597	-0.03349 0.0006 10597	0.02293 0.0362 8416
INCONCUN	0.00705 0.4793 10070	0.00218 0.8269 10070	0.01079 0.2789 10070	-0.05605 0.0001 12356	-0.05411 0.0001 12356	-0.05583 0.0001 12356	-0.01191 0.3284 6733	-0.00512 0.6745 6733	-0.01191 0.3284 6733	-0.03128 0.0021 9660	-0.02801 0.0059 9660	-0.02579 0.0112 9660	-0.01430 0.2110 7653
RACEWIHO	0.04101 0.0001 10993	0.02072 0.0298 10993	0.05381 0.0001 10993	-0.03795 0.0001 13502	-0.04895 0.0001 13502	-0.03022 0.0001 13502	-0.06674 0.0001 7393	-0.03716 0.0014 7393	-0.06674 0.0001 7393	-0.02893 0.0029 10598	-0.01228 0.2061 10598	-0.04040 0.0001 10598	0.01258 0.2500 8366
SEXMIFO	0.00795 0.3992 11241	0.00754 0.4243 11241	0.01124 0.2334 11241	-0.01687 0.0477 13772	-0.01700 0.0460 13772	-0.01704 0.0466 13772	-0.00759 0.5106 7524	-0.00810 0.4825 7524	-0.00759 0.5106 7524	-0.01573 0.1023 10788	-0.01375 0.1532 10788	-0.01847 0.0551 10788	0.00173 0.8728 8555
AGE	-0.00303 0.7479 11241	-0.00331 0.7258 11241	-0.00508 0.5903 11241	0.00614 0.4712 13772	0.00957 0.2616 13772	0.00592 0.4871 13772	0.01721 0.1354 7524	0.01880 0.1029 7524	0.01721 0.1354 7524	0.00684 0.4773 10788	0.01243 0.1967 10788	0.00448 0.6420 10788	-0.00595 0.5819 8555
AGESQ	-0.00215 0.8200 11241	-0.00234 0.8038 11241	-0.00462 0.6244 11241	0.00916 0.2823 13772	0.01297 0.1280 13772	0.00896 0.2932 13772	0.02059 0.0741 7524	0.02363 0.0404 7524	0.02059 0.0741 7524	0.00821 0.3937 10788	0.01371 0.1544 10788	0.00516 0.5922 10788	-0.00280 0.7955 8555

Table 3-6. (continued)

CORRELATION COEFFICIENTS / PROB > |R| UNDER HO:RHO=0 / NUMBER OF OBSERVATIONS

	O3HR01	O3HR03	O3HR04	SPHR01	SPHR03	SPHR04	S4HR01	S4HR03	S4HR04	S2HR01	S2HR03	S2HR04	N2HR01
FAT	-0.00554 0.5623 10959	-0.00429 0.6533 10959	0.00329 0.7304 10959	-0.00336 0.7019 13434	0.00346 0.6883 13434	-0.00336 0.6971 13434	0.01321 0.2576 7345	0.01173 0.3147 7345	0.01321 0.2576 7345	0.00861 0.3773 10513	0.00266 0.7847 10513	0.01301 0.1822 10513	-0.02365 0.0307 8353
FATSU	-0.00674 0.4806 10959	-0.00493 0.6060 10959	0.00165 0.8626 10959	-0.00197 0.8195 13434	0.00517 0.5490 13434	-0.00203 0.8144 13434	0.01287 0.2699 7345	0.01122 0.3365 7345	0.01287 0.2699 7345	0.00893 0.3599 10513	0.00281 0.7736 10513	0.01291 0.1857 10513	-0.02575 0.0186 8353
MARYING	0.02368 0.0121 11241	0.01133 0.2297 11241	0.02628 0.0053 11241	-0.01907 0.0252 13772	-0.02478 0.0036 13772	-0.01902 0.0256 13772	-0.02187 0.0578 7524	-0.02332 0.0431 7524	-0.02187 0.0578 7524	-0.02546 0.0082 10788	-0.01291 0.1800 10788	-0.03675 0.0001 10788	-0.03849 0.0004 8555
VETYING	0.00222 0.8156 11030	-0.00354 0.7103 11030	0.01262 0.1852 11030	-0.02185 0.0111 13514	-0.02006 0.0197 13514	-0.02196 0.0107 13514	-0.03107 0.0076 7385	-0.02999 0.0100 7385	-0.03107 0.0076 7385	-0.02324 0.0169 10572	-0.00910 0.3496 10572	-0.03572 0.0002 10572	-0.02617 0.0165 8400
SMOKYING	-0.01712 0.0787 10548	-0.02309 0.0177 10548	-0.01358 0.1630 10548	0.00932 0.2894 12930	0.01123 0.2017 12930	0.00925 0.2931 12930	0.01353 0.2558 7054	-0.00109 0.9271 7054	0.01353 0.2558 7054	0.03431 0.0005 10147	0.03325 0.0008 10147	0.01752 0.0776 10147	-0.02387 0.0322 8057
NCIGSDYN	-0.01322 0.1793 10318	-0.02214 0.0245 10318	-0.00651 0.5083 10318	-0.00460 0.6049 12638	0.00042 0.9625 12638	-0.00472 0.5960 12638	0.01600 0.1844 6879	0.00786 0.5147 6879	0.01600 0.1844 6879	0.02409 0.0164 9928	0.02804 0.0052 9928	0.01115 0.2664 9928	-0.02512 0.0258 7882
CHRLNDUH	0.01103 0.2423 11241	0.01354 0.1512 11241	0.00321 0.7339 11241	0.03343 0.0001 13772	0.03078 0.0003 13772	0.03320 0.0001 13772	0.01458 0.2061 7524	0.01210 0.2939 7524	0.01458 0.2061 7524	-0.00546 0.5709 10788	0.00293 0.7611 10788	-0.00728 0.4498 10788	0.01598 0.1395 8555
CITYI100	-0.05785 0.0001 11241	-0.03366 0.0004 11241	-0.06421 0.0001 11241	0.08381 0.0001 13772	0.08857 0.0001 13772	0.08370 0.0001 13772	0.02014 0.0807 7524	-0.00544 0.6370 7524	0.02014 0.0807 7524	0.05775 0.0001 10788	0.03141 0.0011 10788	0.08660 0.0001 10788	0.05688 0.0001 8555
PSIYING	0.06091 0.0001 10035	0.09387 0.0001 10035	0.02086 0.0366 10035	0.02149 0.0218 11391	0.02302 0.0140 11391	0.02136 0.0226 11391	0.08202 0.0001 6674	0.12969 0.0001 6674	0.08202 0.0001 6674	-0.00573 0.5740 9635	-0.05603 0.0001 9635	0.07182 0.0001 9635	0.16694 0.0001 7720
BLIN	0.00330 0.7817 7049	0.00964 0.4183 7049	0.00576 0.6287 7049	0.01041 0.3330 8648	0.00426 0.6921 8648	0.01033 0.3370 8648	0.00336 0.8188 4651	-0.00056 0.9697 4651	0.00336 0.8188 4651	0.02010 0.0985 6756	0.01925 0.1136 6756	0.00680 0.5762 6756	-0.03157 0.0210 5343
INDOCCAV	-0.03284 0.0081 6501	-0.02883 0.0201 6501	-0.03635 0.0034 6501	-0.01704 0.1282 7971	-0.00608 0.5873 7971	-0.01705 0.1281 7971	-0.00910 0.5513 4286	-0.01595 0.2965 4286	-0.00910 0.5513 4286	0.00186 0.8834 6238	0.00091 0.9426 6238	0.00613 0.6283 6238	-0.01105 0.4406 4875
CROWDING	0.05429 0.0001 11060	0.06311 0.0001 11060	0.04393 0.0001 11060	0.05502 0.0001 13514	0.04283 0.0001 13514	0.05492 0.0001 13514	0.00947 0.4159 7386	0.01474 0.2051 7386	0.00947 0.4159 7386	0.00253 0.7941 10610	-0.02042 0.0354 10610	0.02006 0.0388 10610	0.06006 0.0001 8405
POLLENHF	0.01310 0.3685 4712	0.01446 0.3211 4712	0.02149 0.1402 4712	-0.01619 0.2547 4952	-0.01963 0.1672 4952	-0.01619 0.2547 4952	0.08192 0.0001 3280	0.03892 0.0258 3280	0.08192 0.0001 3280	-0.05219 0.0004 4581	-0.01588 0.2825 4581	-0.08337 0.0001 4581	-0.14981 0.0001 3715

Table 3-6. (continued)

```
                          CORRELATION COEFFICIENTS / PROB > |R| UNDER HO:RHO=0 / NUMBER OF OBSERVATIONS

            O3NR01    O3NR03    O3NR04    SPNR01    SPNR03    SPNR04    S4NR01    S4NR03    S4NR04    S2NR01    S2NR03    S2NR04    N2NR01

PCDOCH     0.00390   0.05776  -0.06119  -0.05406  -0.05087  -0.05410  -0.04921  -0.00405  -0.04921  -0.06991  -0.10403   0.01400   0.11760
            0.6792    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.7251    0.0001    0.0001    0.0001    0.1459    0.0001
             11241     11241     11241     13772     13772     13772      7524      7524      7524     10788     10788     10788      8555
-------------------------------------------------------------------------------------------------------------------------------------------
            N2NR03    N2NR04    CONR01    CONR03    CONR04  AVMAXTMP  AVMINTMP  DMAXTEMP  DMINTEMP   HUMIDRF  AVPRECIP  EDCOMCON  INCOMCON

O3NR01     0.23570   0.18416  -0.00256  -0.05375   0.05949   0.53117   0.53371   0.51440  -0.37831  -0.13215  -0.18083  -0.00771   0.00705
            0.0001    0.0001    0.7966    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.4177    0.4793
              8168      8168     10100     10100     10100     11241     11241     11241     11241     11152     11241     11050     10070
O3NR03     0.26558   0.23099   0.00477  -0.04083   0.07131   0.52935   0.54489   0.52539  -0.35112  -0.11550  -0.15875  -0.01032   0.00218
            0.0001    0.0001    0.6318    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.2779    0.8269
              8168      8168     10100     10100     10100     11241     11241     11241     11241     11152     11241     11050     10070
O3NR04     0.11077   0.05032  -0.08982  -0.13302  -0.03780   0.49404   0.47866   0.49424  -0.34033  -0.12895  -0.12789  -0.00935   0.01079
            0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.3258    0.2789
              8168      8168     10100     10100     10100     11241     11241     11241     11241     11152     11241     11050     10070
SPNR01     0.35045   0.34742   0.21101   0.16531   0.21081   0.20982   0.15289   0.12991  -0.04644  -0.22564  -0.24960  -0.06004  -0.05605
            0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001
              8535      8535     10934     10934     10934     13739     13739     13739     13739     13673     13739     13542     12356
SPNR03     0.35978   0.36103   0.22987   0.19727   0.22075   0.14624   0.09166   0.06275   0.00125  -0.19222  -0.20708  -0.05074  -0.05411
            0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.8835    0.0001    0.0001    0.0001    0.0001
              8535      8535     10934     10934     10934     13739     13739     13739     13739     13673     13739     13542     12356
SPNR04     0.35036   0.34748   0.21103   0.16517   0.21090   0.21003   0.15279   0.12999  -0.04640  -0.22577  -0.24976  -0.05987  -0.05583
            0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001
              8535      8535     10934     10934     10934     13739     13739     13739     13739     13673     13739     13542     12356
S4NR01     0.01966   0.04659  -0.07036  -0.06502  -0.04153   0.02931   0.05864   0.19349  -0.24462   0.13899   0.09979  -0.02824  -0.01191
            0.1551    0.0007    0.0001    0.0001    0.0007    0.0110    0.0001    0.0001    0.0001    0.0001    0.0001    0.0150    0.3284
              5232      5232      6624      6624      6624      7518      7518      7518      7518      7466      7518      7415      6733
S4NR03     0.16029   0.17785   0.01983   0.01594   0.05323   0.03185   0.06925   0.17368  -0.20745   0.12728   0.11178  -0.03053  -0.00512
            0.0001    0.0001    0.1067    0.1945    0.0001    0.0058    0.0001    0.0001    0.0001    0.0001    0.0001    0.0086    0.6745
              5232      5232      6624      6624      6624      7518      7518      7518      7518      7466      7518      7415      6733
S4NR04     0.01966   0.04659  -0.07036  -0.06502  -0.04153   0.02931   0.05864   0.19349  -0.24462   0.13899   0.09979  -0.02824  -0.01191
            0.1551    0.0007    0.0001    0.0001    0.0007    0.0110    0.0001    0.0001    0.0001    0.0001    0.0001    0.0150    0.3284
              5232      5232      6624      6624      6624      7518      7518      7518      7518      7466      7518      7415      6733
S2NR01    -0.03390  -0.01694  -0.03179  -0.00659  -0.04479  -0.35171  -0.36709  -0.19798  -0.00420   0.07197  -0.06109  -0.03263  -0.03128
            0.0028    0.1348    0.0018    0.5178    0.0001    0.0001    0.0001    0.0001    0.6631    0.0001    0.0001    0.0008    0.0021
              7794      7794      9645      9645      9645     10788     10788     10788     10788     10713     10788     10597      9660
S2NR03    -0.05617  -0.08465  -0.05894  -0.02743  -0.08248  -0.23479  -0.25816  -0.12372  -0.02130   0.05504  -0.01673  -0.02607  -0.02801
            0.0001    0.0001    0.0001    0.0071    0.0001    0.0001    0.0001    0.0001    0.0269    0.0001    0.0822    0.0073    0.0059
              7794      7794      9645      9645      9645     10788     10788     10788     10788     10713     10788     10597      9660
S2NR04    -0.02850   0.03985  -0.02071  -0.00751  -0.01063  -0.45246  -0.43773  -0.24764   0.00210   0.12062  -0.05051  -0.03349  -0.02579
            0.0119    0.0004    0.0420    0.4607    0.2967    0.0001    0.0001    0.0001    0.8271    0.0001    0.0001    0.0006    0.0112
              7794      7794      9645      9645      9645     10788     10788     10788     10788     10713     10788     10597      9660
```

Table 3-6. (continued)

CORRELATION COEFFICIENTS / PROB > |R| UNDER HO:RHO=0 / NUMBER OF OBSERVATIONS

	N2NR03	N2NR04	CONR01	CONR03	CONR04	AVMAXTMP	AVMINTMP	DMAXTEMP	DMINTEMP	HUMIDRF	AVPRECIP	EDCONCON	INCONCON
N2NR01	0.90504 0.0001 8555	0.95640 0.0001 8555	0.39543 0.0001 8287	0.29751 0.0001 8287	0.43521 0.0001 8287	0.08481 0.0001 8555	0.10962 0.0001 8555	-0.03787 0.0005 8555	0.12348 0.0001 8555	-0.01544 0.1552 6490	-0.22783 0.0001 8555	0.02283 0.0362 8416	-0.01430 0.2110 7653
N2NR03	1.00000 0.0000 8555	0.83720 0.0001 8555	0.39474 0.0001 8287	0.32599 0.0001 8287	0.42723 0.0001 8287	0.04870 0.0001 8555	0.07833 0.0001 8555	-0.07015 0.0001 8555	0.14617 0.0001 8555	0.00604 0.5779 8480	-0.18553 0.0001 8555	0.01526 0.1615 8416	-0.01969 0.0850 7653
N2NR04	0.83720 0.0001 8555	1.00000 0.0000 8555	0.37241 0.0001 8287	0.27632 0.0001 8287	0.42771 0.0001 8287	-0.01078 0.3187 8555	0.02898 0.0073 8555	-0.10232 0.0001 8555	0.15262 0.0001 8555	0.03550 0.0011 8480	-0.20088 0.0001 8555	0.01794 0.0998 8416	-0.02357 0.0392 7653
CONR01	0.39479 0.0001 8287	0.37241 0.0001 8287	1.00000 0.0000 10962	0.92735 0.0001 10962	0.95091 0.0001 10962	-0.06294 0.0001 10929	-0.05276 0.0001 10929	-0.13912 0.0001 10929	0.14751 0.0001 10929	-0.08533 0.0001 10878	-0.06815 0.0001 10929	0.04662 0.0001 10772	-0.00653 0.5175 9809
CONR03	0.32599 0.0001 8287	0.27632 0.0001 8287	0.92735 0.0001 10962	1.00000 0.0000 10962	0.83416 0.0001 10962	-0.11406 0.0001 10929	-0.10948 0.0001 10929	-0.18799 0.0001 10929	0.17267 0.0001 10929	-0.06651 0.0001 10878	-0.04163 0.0001 10929	0.03114 0.0012 10772	-0.01375 0.1733 9809
CONR04	0.42723 0.0001 8287	0.42771 0.0001 8287	0.95091 0.0001 10962	0.83416 0.0001 10962	1.00000 0.0000 10962	-0.05204 0.0001 10929	-0.02345 0.0142 10929	-0.10553 0.0001 10929	0.11919 0.0001 10929	-0.04008 0.0001 10878	-0.07803 0.0001 10929	0.04173 0.0001 10772	-0.01152 0.2541 9809
AVMAXTMP	0.04870 0.0001 8555	-0.01078 0.3187 8555	-0.06294 0.0001 10929	-0.11406 0.0001 10929	-0.05204 0.0001 10929	1.00000 0.0000 14155	0.96563 0.0001 14155	0.71762 0.0001 14155	-0.24008 0.0001 14155	-0.35946 0.0001 14053	0.00953 0.2571 14155	-0.02326 0.0061 13920	-0.00985 0.2669 12698
AVMINTMP	0.07833 0.0001 8555	0.02898 0.0073 8555	-0.05276 0.0001 10929	-0.10948 0.0001 10929	-0.02345 0.0142 10929	0.96563 0.0001 14155	1.00000 0.0000 14155	0.71773 0.0001 14155	-0.23723 0.0001 14155	-0.22815 0.0001 14053	0.07269 0.0001 14155	-0.01701 0.0447 13920	-0.00320 0.7183 12698
DMAXTEMP	-0.07015 0.0001 8555	-0.10232 0.0001 8555	-0.13912 0.0001 10929	-0.18799 0.0001 10929	-0.10553 0.0001 10929	0.71762 0.0001 14155	0.71773 0.0001 14155	1.00000 0.0000 14155	-0.83984 0.0001 14155	-0.17158 0.0001 14053	0.04233 0.0001 14155	-0.01686 0.0467 13920	0.01968 0.0266 12698
DMINTEMP	0.14617 0.0001 8555	0.15262 0.0001 8555	0.14751 0.0001 10929	0.17267 0.0001 10929	0.11919 0.0001 10929	-0.24008 0.0001 14155	-0.23723 0.0001 14155	-0.83984 0.0001 14155	1.00000 0.0000 14155	0.01527 0.0702 14053	-0.02854 0.0007 14155	0.01051 0.2150 13920	-0.03000 0.0007 12698
HUMIDRF	0.00604 0.5779 8480	0.03550 0.0011 8480	-0.08533 0.0001 10878	-0.06651 0.0001 10878	-0.04008 0.0001 10878	-0.35946 0.0001 14053	-0.22815 0.0001 14053	-0.17158 0.0001 14053	0.01527 0.0702 14053	1.00000 0.0000 14088	0.20355 0.0001 14053	0.01605 0.0588 13856	0.01531 0.0852 12634
AVPRECIP	-0.18553 0.0001 8555	-0.20088 0.0001 8555	-0.06815 0.0001 10929	-0.04163 0.0001 10929	-0.07803 0.0001 10929	0.00953 0.2571 14155	0.07269 0.0001 14155	0.04233 0.0001 14155	-0.02854 0.0007 14155	0.20355 0.0001 14053	1.00000 0.0000 14155	-0.00701 0.4086 13920	-0.00507 0.5678 12698
EDCONCON	0.01526 0.1615 8416	0.01794 0.0998 8416	0.04662 0.0001 10772	0.03114 0.0012 10772	0.04173 0.0001 10772	-0.02326 0.0061 13920	-0.01701 0.0447 13920	-0.01686 0.0467 13920	0.01051 0.2150 13920	0.01605 0.0588 13856	-0.00701 0.4086 13920	1.00000 0.0000 14204	0.33807 0.0001 12853

Table 3-6. (continued)

```
                      CORRELATION COEFFICIENTS / PROB > |R| UNDER HO:RHO=0 / NUMBER OF OBSERVATIONS

           U2NR03   U2NR04   CONR01   CONR03   CONR04 AVMAXTMP AVMINTMP DMAXTEMP DMINTEMP  HUMIDRF AVPRECIP EDCOMCON INCOMCON

INCOMCON -0.01969 -0.02357 -0.00653 -0.01375 -0.01152 -0.00985 -0.00320  0.01968 -0.03000  0.01531 -0.00507  0.33807  1.00000
           0.0850   0.0392   0.5175   0.1733   0.2541   0.2669   0.7183   0.0266   0.0007   0.0852   0.5678   0.0001   0.0000
             7653     7653     9809     9809     9809    12698    12698    12698    12698    12634    12698    12853    12963

RACE*100  0.00132 -0.00809  0.04805  0.03461  0.04913 -0.03297 -0.04001 -0.01477 -0.00298  0.01166 -0.04688  0.12117  0.18013
           0.9036   0.4596   0.0001   0.0003   0.0001   0.0001   0.0001   0.0819   0.7259   0.1706   0.0001   0.0001   0.0001
             8366     8366    10721    10721    10721    13878    13878    13878    13878    13811    13878    13925    12715

SEXMIFG   0.00114 -0.00020 -0.00387 -0.00416 -0.00695 -0.00625 -0.00707 -0.01271  0.01123  0.01730  0.00156  0.04838  0.09687
           0.9163   0.9856   0.6857   0.6630   0.4669   0.4573   0.4000   0.1306   0.1816   0.0401   0.8531   0.0001   0.0001
             8555     8555    10962    10962    10962    14155    14155    14155    14155    14088    14155    14204    12963

AGE      -0.00185 -0.00487 -0.00825  0.00025 -0.00958  0.01186  0.01629  0.00580  0.00482 -0.00459  0.01419 -0.24765 -0.15482
           0.8642   0.6527   0.3880   0.9789   0.3161   0.1584   0.0526   0.4902   0.5663   0.5859   0.0915   0.0001   0.0001
             8555     8555    10962    10962    10962    14155    14155    14155    14155    14088    14155    14204    12963

AGESQ     0.00152 -0.00245 -0.00481  0.00346 -0.00714  0.01215  0.01569  0.00535  0.00517 -0.00470  0.01160 -0.27678 -0.19773
           0.8882   0.8210   0.6146   0.7171   0.4649   0.1484   0.0622   0.5243   0.5387   0.5771   0.1675   0.0001   0.0001
             8555     8555    10962    10962    10962    14155    14155    14155    14155    14088    14155    14204    12963

FAT      -0.02654 -0.01824 -0.03640 -0.03134 -0.04002 -0.01888 -0.02147 -0.00550 -0.00843  0.01700  0.01166 -0.07883  0.00913
           0.0153   0.0956   0.0002   0.0012   0.0001   0.0265   0.0116   0.5180   0.3218   0.0462   0.1706   0.0001   0.3024
             8353     8353    10691    10691    10691    13811    13811    13811    13811    13745    13811    13985    12752

FATSQ    -0.02905 -0.02019 -0.03630 -0.03145 -0.03966 -0.01841 -0.02180 -0.00475 -0.00949  0.01599  0.01944 -0.08080  0.00212
           0.0079   0.0651   0.0002   0.0011   0.0001   0.0305   0.0104   0.5769   0.2649   0.0609   0.2197   0.0001   0.8109
             8353     8353    10691    10691    10691    13811    13811    13811    13811    13745    13811    13985    12752

MARYIND  -0.03565 -0.04986 -0.01966 -0.01565 -0.02279  0.01735  0.01160  0.00804  0.00149 -0.00275  0.00419  0.05492  0.25007
           0.0010   0.0001   0.0395   0.1012   0.0170   0.0390   0.1676   0.3391   0.8597   0.7437   0.6183   0.0001   0.0001
             8555     8555    10962    10962    10962    14155    14155    14155    14155    14088    14155    14204    12963

VETYIND  -0.02186 -0.03225 -0.01956 -0.01661 -0.02713  0.00353 -0.00164 -0.00104  0.00230 -0.00293 -0.00237  0.08383  0.12891
           0.0451   0.0031   0.0426   0.0850   0.0049   0.6776   0.8471   0.9023   0.7860   0.7303   0.7796   0.0001   0.0001
             8400     8400    10751    10751    10751    13893    13893    13893    13893    13826    13893    14035    12789

SMOKYIND -0.01406 -0.02122 -0.00683 -0.00638 -0.00125  0.00100 -0.00569  0.00266 -0.00709 -0.00689  0.00128 -0.03087 -0.02302
           0.2071   0.0568   0.4884   0.5174   0.8987   0.9079   0.5118   0.7595   0.4142   0.4282   0.8825   0.0004   0.0107
             8057     8057    10304    10304    10304    13280    13280    13280    13280    13214    13280    13389    12284

NCIGSDYD -0.01774 -0.02510 -0.02361 -0.01954 -0.02363  0.00534  0.00192  0.01175 -0.01586 -0.00302  0.01549 -0.01844  0.00633
           0.1152   0.0259   0.0177   0.0497   0.0176   0.5428   0.8269   0.1810   0.0709   0.7314   0.0777   0.0349   0.4879
             7882     7882    10086    10086    10086    12972    12972    12972    12972    12905    12972    13077    11999

CHILDNUM  0.01318  0.01785 -0.00866 -0.00545 -0.01013  0.01720  0.01435  0.00835  0.00011 -0.00050  0.00064 -0.21406 -0.20345
           0.2228   0.0988   0.3648   0.5682   0.2888   0.0408   0.0878   0.3206   0.9892   0.9525   0.9393   0.0001   0.0001
             8555     8555    10962    10962    10962    14155    14155    14155    14158    14088    14155    14204    12963

CITYI100  0.04458  0.06459  0.02852  0.03270  0.01667  0.00500 -0.00394  0.00460 -0.01110 -0.05539 -0.03722 -0.09389 -0.16819
           0.0001   0.0001   0.0028   0.0006   0.0810   0.5522   0.6390   0.5844   0.1867   0.0001   0.0001   0.0001   0.0001
             8555     8555    10962    10962    10962    14155    14155    14155    14155    14088    14155    14204    12963
```

Table 3-6. (continued)

CORRELATION COEFFICIENTS / PROB > |R| UNDER HO:RHO=0 / NUMBER OF OBSERVATIONS

	N2NN03	N2NN04	CONN01	CONN03	CONN04	AVMAXTMP	AVMINTMP	DMAXTEMP	DMINTEMP	HUMIDAF	AVPRECIP	EDCOMCON	INCOMCON
PSIYIND	0.14686 0.0001 7726	0.20914 0.0001 7720	0.04626 0.0001 10018	0.05465 0.0001 10018	0.06452 0.0001 10018	-0.03710 0.0001 11624	0.00203 0.8268 11624	0.00696 0.4533 11624	-0.02098 0.0237 11624	-0.06414 0.0001 11591	-0.00900 0.3319 11624	-0.01415 0.1267 11651	-0.00701 0.4707 10600
BLUF	-0.01866 0.1726 5343	-0.03069 0.0249 5343	-0.03138 0.0092 6880	-0.02541 0.0351 6880	-0.03205 0.0078 6880	-0.00523 0.6222 8873	-0.00898 0.3974 8873	0.00617 0.5609 8873	-0.01327 0.2112 8873	0.02005 0.0595 8833	0.00545 0.6081 8873	-0.43064 0.0001 9915	-0.18615 0.0001 8335
INDOCCAV	-0.01507 0.2927 4875	-0.00707 0.6218 4875	-0.00662 0.5983 6339	0.00425 0.7353 6339	-0.01720 0.1709 6339	0.00610 0.5814 8167	0.00789 0.4759 8167	0.00108 0.9219 8167	0.00423 0.7025 8167	0.00259 0.8154 8132	0.00018 0.9870 8167	0.32669 0.0001 8297	0.12508 0.0001 7684
CPOGGIND	0.05972 0.0001 8405	0.05807 0.0001 8405	0.00375 0.6968 10786	-0.00027 0.9777 10786	0.01396 0.1471 10786	0.04452 0.0001 13890	0.04865 0.0001 13890	0.03177 0.0002 13890	-0.00880 0.2998 13890	-0.01683 0.0478 13823	-0.01186 0.1623 13890	-0.14673 0.0001 13966	0.00967 0.2747 12763
POLLEINF	-0.14622 0.0001 3715	-0.15632 0.0001 3715	-0.12838 0.0001 4503	-0.12647 0.0001 4503	-0.13760 0.0001 4503	0.20438 0.0001 5014	0.20442 0.0001 5014	0.22360 0.0001 5014	-0.16761 0.0001 5014	0.05522 0.0001 4983	0.20808 0.0001 5014	-0.00525 0.7127 4929	-0.00095 0.9495 4460
PCDOCCU	0.12686 0.0001 8555	0.12957 0.0001 8555	0.22128 0.0001 10962	0.19589 0.0001 10962	0.24372 0.0001 10962	0.00304 0.7173 14155	0.06666 0.0001 14155	0.00484 0.5648 14155	0.02575 0.0022 14155	-0.01656 0.0494 14088	0.00316 0.7074 14155	0.05655 0.0001 14204	0.02330 0.0080 12963

	RACE-IND	SEXMIFO	AGE	AGESQ	FAT	FATSQ	MARYIND	VETYIND	SMOKYIND	NCIGSDYM	CHRLMDUM	CITYI100	PSIYIND
O3NN01	0.04101 0.0001 10993	0.00795 0.3992 11241	-0.00303 0.7479 11241	-0.00215 0.8200 11241	-0.00554 0.5623 10959	-0.00674 0.4806 10959	0.02368 0.0121 11241	0.00222 0.8156 11030	-0.01712 0.0767 10548	-0.01322 0.1793 10318	0.01103 0.2423 11241	-0.05785 0.0001 11241	0.06091 0.0001 10035
O3NN03	0.02072 0.0298 10993	0.00754 0.4243 11241	-0.00331 0.7256 11241	-0.00234 0.8038 11241	-0.00429 0.6533 10959	-0.00493 0.6060 10959	0.01133 0.2297 11241	-0.00354 0.7103 11030	-0.02309 0.0177 10548	-0.02214 0.0245 10318	0.01354 0.1512 11241	-0.03366 0.0004 11241	0.09387 0.0001 10035
O3NN04	0.05381 0.0001 10993	0.01124 0.2334 11241	-0.00508 0.5903 11241	-0.00462 0.6244 11241	0.00329 0.7309 10959	0.00165 0.8626 10959	0.02628 0.0053 11241	0.01262 0.1852 11030	-0.01358 0.1630 10548	-0.00651 0.5083 10318	0.00321 0.7339 11241	-0.06421 0.0001 11241	0.02086 0.0366 10035
SPNN01	-0.03796 0.0001 13502	-0.01687 0.0477 13772	0.00614 0.4712 13772	0.00916 0.2823 13772	-0.00330 0.7019 13434	-0.00197 0.8195 13434	-0.01907 0.0252 13772	-0.02185 0.0111 13514	0.00932 0.2894 12930	-0.00460 0.6049 12638	0.03343 0.0001 13772	0.08381 0.0001 13772	0.02149 0.0218 11391
SPNN03	-0.04895 0.0001 13502	-0.01700 0.0460 13772	0.00957 0.2616 13772	0.01297 0.1280 13772	0.00346 0.6883 13434	0.00517 0.5490 13434	-0.02478 0.0036 13772	-0.02006 0.0197 13514	0.01123 0.2017 12930	0.00042 0.9625 12638	0.03078 0.0003 13772	0.08857 0.0001 13772	0.02302 0.0140 11391
SPNN04	-0.03822 0.0001 13502	-0.01704 0.0456 13772	0.00592 0.4871 13772	0.00896 0.2932 13772	-0.00336 0.6971 13434	-0.00203 0.8144 13434	-0.01902 0.0256 13772	-0.02196 0.0107 13514	0.00925 0.2931 12930	-0.00472 0.5960 12638	0.03320 0.0001 13772	0.08370 0.0001 13772	0.02136 0.0226 11391
S4NN01	-0.06674 0.0001 7193	-0.00759 0.5106 7524	0.01721 0.1354 7524	0.02059 0.0741 7524	0.01321 0.2576 7345	0.01287 0.2697 7364	-0.02187 0.0578 7524	-0.03107 0.0076 7385	0.01353 0.2558 7054	0.01600 0.1844 6879	0.01458 0.2061 7524	0.02014 0.0807 7524	0.08202 0.0001 6674

Table 3-6. (continued)

CORRELATION COEFFICIENTS / PROB > |R| UNDER HO:RHO=0 / NUMBER OF OBSERVATIONS

	RACEWIND	SEXMIFO	AGE	AGESQ	FAT	FATSQ	MARYIND	VETYIND	SMOKYIND	NCIGSDYN	CHALMDUM	CITYI100	PSIYIND
S4NR03	-0.03716 0.0014 7393	-0.00810 0.4825 7524	0.01880 0.1029 7524	0.02363 0.0404 7524	0.01171 0.3147 7345	0.01122 0.3365 7345	-0.02332 0.0431 7524	-0.02995 0.0100 7385	-0.00109 0.9271 7054	0.00786 0.5147 6879	0.01210 0.2939 7524	-0.00544 0.6370 7524	0.12969 0.0001 6674
S4NR04	-0.04674 0.0001 7393	-0.00759 0.5106 7524	0.01721 0.1354 7524	0.02059 0.0741 7524	0.01321 0.2576 7345	0.01287 0.2699 7345	-0.02187 0.0578 7524	-0.03107 0.0076 7385	0.01353 0.2558 7054	0.01600 0.1844 6879	0.01458 0.2061 7524	0.02014 0.0807 7524	0.08202 0.0001 6674
S2NR01	-0.02893 0.0029 10598	-0.01573 0.1023 10788	0.00684 0.4773 10788	0.00821 0.3937 10788	0.00861 0.3773 10513	0.00893 0.3599 10513	-0.02546 0.0082 10788	-0.02324 0.0169 10572	0.03431 0.0006 10147	0.02409 0.0164 9928	-0.00546 0.5709 10788	0.05775 0.0001 10788	-0.00573 0.5740 9635
S2NR03	-0.01228 0.2061 10598	-0.01375 0.1532 10788	0.01243 0.1967 10788	0.01371 0.1544 10788	0.00266 0.7867 10513	0.00281 0.7736 10513	-0.01291 0.1800 10788	-0.00910 0.3496 10572	0.03325 0.0008 10147	0.02804 0.0052 9928	0.00293 0.7611 10788	0.03141 0.0011 10788	-0.05603 0.0001 9635
S2NR04	-0.04040 0.0001 10598	-0.01847 0.0551 10788	0.00448 0.6420 10788	0.00516 0.5922 10788	0.01301 0.1822 10513	0.01291 0.1857 10513	-0.03675 0.0001 10788	-0.03572 0.0002 10572	0.01752 0.0776 10147	0.01115 0.2664 9928	-0.00728 0.4498 10788	0.08660 0.0001 10788	0.07182 0.0001 9635
N2NR01	0.01258 0.2500 8366	0.00173 0.8728 8555	-0.00595 0.5819 8555	-0.00280 0.7955 8555	-0.02365 0.0307 8353	-0.02575 0.0186 8353	-0.03849 0.0004 8555	-0.02617 0.0165 8400	-0.02387 0.0322 8057	-0.02512 0.0258 7882	0.01598 0.1395 8555	0.05688 0.0001 8555	0.16694 0.0001 7720
N2NR03	0.00132 0.9036 8366	0.00114 0.9163 8555	-0.00185 0.8642 8555	0.00152 0.8882 8555	-0.02654 0.0153 8353	-0.02905 0.0079 8353	-0.03565 0.0010 8555	-0.02186 0.0451 8400	-0.01406 0.2071 8057	-0.01774 0.1152 7882	0.01318 0.2228 8555	0.04858 0.0001 8555	0.14688 0.0001 7720
N2NR04	-0.00809 0.4596 8366	-0.00020 0.9856 8555	-0.00487 0.6527 8555	-0.00245 0.8210 8555	-0.01824 0.0956 8353	-0.02019 0.0651 8353	-0.04986 0.0001 8555	-0.03225 0.0031 8400	-0.02122 0.0568 8057	-0.02510 0.0259 7882	0.01785 0.0988 8555	0.06459 0.0001 8555	0.20914 0.0001 7720
CONR01	0.04805 0.0001 10721	-0.00387 0.6857 10962	-0.00825 0.3880 10962	-0.00481 0.6146 10962	-0.03640 0.0002 10691	-0.03630 0.0002 10691	-0.01966 0.0395 10962	-0.01956 0.0426 10751	-0.00683 0.4884 10304	-0.02361 0.0177 10086	-0.00866 0.3648 10962	0.02852 0.0028 10962	0.04626 0.0001 10018
CONR03	0.03461 0.0003 10721	-0.00416 0.6630 10962	0.00025 0.9789 10962	0.00346 0.7171 10962	-0.03134 0.0012 10691	-0.03145 0.0011 10691	-0.01565 0.1012 10962	-0.01661 0.0650 10751	-0.00638 0.5174 10304	-0.01954 0.0497 10086	-0.00545 0.5682 10962	0.03270 0.0006 10962	0.05465 0.0001 10018
CONR04	0.04913 0.0001 10721	-0.00695 0.4669 10962	-0.00958 0.3161 10962	-0.00714 0.4549 10962	-0.04002 0.0001 10691	-0.03966 0.0001 10691	-0.02279 0.0170 10962	-0.02713 0.0049 10751	-0.00125 0.8987 10304	-0.02363 0.0176 10086	-0.01013 0.2888 10962	0.01667 0.0810 10962	0.06452 0.0001 10018
AVMAXTMP	-0.03297 0.0001 13878	-0.00625 0.4573 14155	0.01186 0.1584 14155	0.01215 0.1484 14155	-0.01888 0.0265 13811	-0.01841 0.0305 13811	0.01735 0.0390 14155	0.00353 0.6776 13893	0.00100 0.9079 13280	0.00534 0.5428 12972	0.01720 0.0408 14155	0.00500 0.5522 14155	-0.03718 0.0001 11624
AVMINTMP	-0.04001 0.0001 13878	-0.00707 0.4000 14155	0.01629 0.0526 14155	0.01568 0.0622 14155	-0.02147 0.0116 13811	-0.02180 0.0104 13811	0.01160 0.1676 14155	-0.00164 0.8471 13893	-0.00569 0.5118 13280	0.00192 0.8269 12972	0.01435 0.0878 14155	-0.00394 0.6390 14155	0.00203 0.8268 11624

Table 3-6. (continued)

CORRELATION COEFFICIENTS / PROB > |R| UNDER HO:RHO=0 / NUMBER OF OBSERVATIONS

	RACEWIND	SEXMIFO	AGE	AGESQ	FAT	FATSQ	MARYIND	VETYIND	SMOKYIND	NCIGSDYN	CHRLMDUM	CITYI100	PSIYIND
DMAXILOP	-0.01477 0.0819 13878	-0.01271 0.1306 14155	0.00580 0.4902 14155	0.00535 0.5243 14155	-0.00550 0.5180 13811	-0.00475 0.5769 13811	0.00804 0.3391 14155	-0.00104 0.9023 13893	0.00266 0.7595 13280	0.01175 0.1810 12972	0.00835 0.3206 14155	0.00460 0.5844 14155	0.00696 0.4533 11626
DMINTFOP	-0.00298 0.7259 13878	0.01123 0.1816 14155	0.00482 0.5663 14155	0.00517 0.5387 14155	-0.00843 0.3218 13811	-0.00949 0.2649 13811	0.00149 0.8597 14155	0.00230 0.7860 13893	-0.00709 0.4142 13280	-0.01586 0.0709 12972	0.00011 0.9892 14155	-0.01110 0.1867 14155	-0.02098 0.0237 11624
HUMIDRF	0.01165 0.1706 13811	0.01730 0.0401 14088	-0.00459 0.5859 14088	-0.00470 0.5771 14088	0.01700 0.0462 13745	0.01599 0.0609 13745	-0.00275 0.7437 14088	-0.00293 0.7303 13826	-0.00689 0.4282 13214	-0.00302 0.7314 12905	-0.00050 0.9525 14088	-0.05535 0.0001 14088	-0.06414 0.0001 11591
AVPRECIP	-0.04688 0.0001 13878	0.00156 0.8531 14155	0.01419 0.0915 14155	0.01160 0.1675 14155	0.01166 0.1706 13811	0.01044 0.2197 13811	0.00419 0.6183 14155	-0.00237 0.7796 13893	0.00128 0.8825 13280	0.01549 0.0777 12972	0.00064 0.9393 14155	-0.03722 0.0001 14155	-0.00900 0.3319 11624
EDCOMCOL	0.12117 0.0001 13925	0.04838 0.0001 14204	-0.24765 0.0001 14204	-0.27678 0.0001 14204	-0.07883 0.0001 13985	-0.08080 0.0001 13985	0.05492 0.0001 14204	0.08383 0.0001 14035	-0.03087 0.0004 13389	-0.01844 0.0349 13077	-0.21406 0.0001 14204	-0.09389 0.0001 14204	-0.01415 0.1267 11651
INCOMCOL	0.16013 0.0001 12715	0.09687 0.0001 12963	-0.15482 0.0001 12963	-0.19773 0.0001 12963	0.00913 0.3024 12752	0.00212 0.8109 12752	0.25007 0.0001 12963	0.12591 0.0001 12789	-0.02302 0.0107 12284	0.00633 0.4879 11999	-0.20345 0.0001 12963	-0.16819 0.0001 12963	-0.00701 0.4707 10600
RACEWIND	1.00000 0.0000 14155	0.02778 0.0009 14155	0.04963 0.0001 14155	0.04767 0.0001 14155	-0.06648 0.0001 13819	-0.06620 0.0001 13819	0.13015 0.0001 14155	0.05048 0.0001 13899	-0.02634 0.0024 13272	0.05079 0.0001 12958	-0.03888 0.0001 14155	-0.24836 0.0001 14155	-0.00040 0.9657 11594
SEXMIFO	0.02778 0.0009 14155	1.00000 0.0000 14441	-0.02696 0.0012 14441	-0.03027 0.0003 14441	0.34606 0.0001 14095	0.30760 0.0001 14095	0.08268 0.0001 14441	0.47976 0.0001 14174	0.06432 0.0001 13547	0.09361 0.0001 13231	-0.00207 0.8031 14441	-0.02109 0.0113 14441	-0.01222 0.1833 11851
AGE	0.04963 0.0001 14155	-0.02696 0.0012 14441	1.00000 0.0000 14441	0.98157 0.0001 14441	0.13866 0.0001 14095	0.12627 0.0001 14095	0.19555 0.0001 14441	0.14028 0.0001 14174	-0.11651 0.0001 13547	-0.07507 0.0001 13231	0.33254 0.0001 14441	0.00782 0.3473 14441	0.03971 0.0001 11851
AGESQ	0.04767 0.0001 14155	-0.03027 0.0003 14441	0.98157 0.0001 14441	1.00000 0.0000 14441	0.09759 0.0001 14095	0.08747 0.0001 14095	0.11758 0.0001 14441	0.09533 0.0001 14174	-0.14097 0.0001 13547	-0.10558 0.0001 13231	0.34372 0.0001 14441	0.01456 0.0802 14441	0.03947 0.0001 11851
FAT	-0.06648 0.0001 13819	0.34606 0.0001 14095	0.13866 0.0001 14095	0.09759 0.0001 14095	1.00000 0.0000 14095	0.98925 0.0001 14095	0.14071 0.0001 14095	0.24205 0.0001 13927	-0.04202 0.0001 13275	-0.00061 0.9444 12965	0.07784 0.0001 14095	0.00979 0.2452 14095	0.00332 0.7211 11562
FATSQ	-0.06620 0.0001 13819	0.30760 0.0001 14095	0.12627 0.0001 14095	0.08747 0.0001 14095	0.98925 0.0001 14095	1.00000 0.0000 14095	0.12877 0.0001 14095	0.22206 0.0001 13927	-0.04065 0.0001 13275	-0.00235 0.7891 12965	0.08145 0.0001 14095	0.01077 0.2010 14095	0.00366 0.6941 11562
MARYIND	0.13015 0.0001 14155	0.08268 0.0001 14441	0.19555 0.0001 14441	0.11758 0.0001 14441	0.14071 0.0001 14095	0.12877 0.0001 14095	1.00000 0.0000 14441	0.17867 0.0001 14174	-0.01175 0.1715 13547	0.02558 0.0033 13231	-0.02563 0.0021 14441	-0.10825 0.0001 14441	-0.02152 0.0191 11851

Table 3-6. (continued)

```
                         CORRELATION COEFFICIENTS / PROB > |R| UNDER HO:RHO=0 / NUMBER OF OBSERVATIONS

          RACEWINO  SEXMIFO      AGE    AGESQ      FAT    FATSQ  MARYINO  VETYINO SMOKYINO NCIGSDYN CHRLMDUM CITY1100  PSIYINO

VETYINO    0.65048  0.47976  0.14028  0.09533  0.24205  0.22206  0.17867  1.00000  0.07133  0.11991  0.02021 -0.04234 -0.01647
            0.0001   0.0001   0.0001   0.0001   0.0001   0.0001   0.0001   0.0000   0.0001   0.0001   0.0161   0.0001   0.0757
             13899    14174    14174    14174    13927    13927    14174    14174    13345    13032    14174    14174    11629

SMOKYINO  -0.02634  0.06432 -0.11651 -0.14097 -0.04202 -0.04065 -0.01175  0.07133  1.00000  0.79205 -0.02155  0.02921  0.00362
            0.0024   0.0001   0.0001   0.0001   0.0001   0.0001   0.1715   0.0001   0.0000   0.0001   0.0121   0.0007   0.7030
             13272    13547    13547    13547    13275    13275    13547    13345    13547    13231    13547    13547    11100

NCIGSDYN   0.05079  0.09361 -0.07507 -0.10558 -0.00061 -0.00235  0.02558  0.11991  0.79205  1.00000 -0.00337  0.00730  0.01066
            0.0001   0.0001   0.0001   0.0001   0.9444   0.7891   0.0033   0.0001   0.0001   0.0000   0.6985   0.4009   0.2666
             12958    13231    13231    13231    12965    12965    13231    13032    13231    13231    13231    13231    10851

CHRLMDUM  -0.03848 -0.00207  0.33254  0.34372  0.07784  0.08145 -0.02563  0.02021 -0.02155 -0.00337  1.00000  0.03161  0.01311
            0.0001   0.8031   0.0001   0.0001   0.0001   0.0001   0.0021   0.0161   0.0121   0.6985   0.0000   0.0001   0.1536
             14155    14441    14441    14441    14095    14095    14441    14174    13547    13231    14441    14441    11851

CITY1100  -0.24836 -0.02109  0.00782  0.01456  0.00979  0.01077 -0.10825 -0.04234  0.02921  0.00730  0.03161  1.00000  0.00742
            0.0001   0.0113   0.3473   0.0802   0.2452   0.2010   0.0001   0.0001   0.0007   0.4009   0.0001   0.0000   0.4193
             14155    14441    14441    14441    14095    14095    14441    14174    13547    13231    14441    14441    11851

PSIYINO   -0.00040 -0.01222  0.03471  0.03947  0.00332  0.00366 -0.02152 -0.01647  0.00362  0.01066  0.01311  0.00742  1.00000
            0.9657   0.1833   0.0001   0.0001   0.7211   0.6941   0.0191   0.0757   0.7030   0.2666   0.1536   0.4193   0.0000
             11594    11851    11851    11851    11562    11562    11851    11629    11100    10851    11851    11851    11851

BLUE      -0.10162  0.22899 -0.06926 -0.05828  0.14127  0.13269 -0.01335  0.04365  0.08951  0.07004  0.00574  0.00903 -0.01183
            0.0001   0.0001   0.0001   0.0001   0.0001   0.0001   0.2038   0.0001   0.0001   0.0001   0.5847   0.3897   0.3074
              8896     9066     9066     9066     8937     8937     9066     8969     8431     8238     9066     9066     7446

INDOCCAV  -0.03721 -0.04156  0.11251  0.09217 -0.01816 -0.01778  0.06132  0.05375 -0.08224 -0.08018  0.01078  0.02283  0.01044
            0.0008   0.0001   0.0001   0.0001   0.0997   0.1070   0.0001   0.0001   0.0001   0.0001   0.3248   0.0370   0.3872
              8190     8345     8345     8345     8217     8217     8345     8253     7768     7593     8345     8345     6863

CROWDING  -0.12859  0.01707 -0.32765 -0.32056 -0.00308  0.00153  0.04870 -0.05664  0.04113  0.01801 -0.09239  0.02672  0.00546
            0.0001   0.0421   0.0001   0.0001   0.7168   0.8572   0.0001   0.0001   0.0001   0.0400   0.0001   0.0015   0.5557
             13888    14171    14171    14171    13854    13854    14171    13933    13313    13003    14171    14171    11649

POLLENHF  -0.04812  0.00652 -0.00033  0.00066  0.02564  0.02762  0.01027  0.00455  0.01323  0.00718  0.00035  0.01368 -0.01681
            0.0007   0.6444   0.9816   0.9626   0.0729   0.0534   0.4672   0.7494   0.3631   0.6256   0.9805   0.3327   0.2369
              4905     5014     5014     5014     4892     4892     5014     4924     4726     4626     5014     5014     4954

PCDOCN    -0.02885  0.00264  0.00232  0.00082 -0.03126 -0.03073 -0.03591 -0.01027 -0.00747 -0.01080 -0.01544  0.04954  0.15820
            0.0006   0.7511   0.7801   0.9214   0.0002   0.0003   0.0001   0.2214   0.3844   0.2143   0.0636   0.0001   0.0001
             14155    14441    14441    14441    14095    14095    14441    14174    13547    13231    14441    14441    11851

              BLUE INDOCCAV CROWDING POLLENHF   PCDOCN

O3HR01     0.00330 -0.03284  0.05429  0.01310  0.00390
            0.7817   0.0081   0.0001   0.3685   0.6792
              7049     6591    11060     4712    11241

O3HR03     0.00966 -0.02883  0.06311  0.01446  0.05776
            0.4183   0.0201   0.0001   0.3211   0.0001
              7049     6591    11060     4712    11241
```

Table 3-6. (continued)

```
                 CORRELATION COEFFICIENTS / PROB > |R| UNDER HO:RHO=0 / NUMBER OF OBSERVATIONS

              BLUE  INDOCCAV CHOWPING POLLENHF    PCOUCH

U3HR04     0.00576 -0.03635  0.04393  0.02149 -0.06119
            0.6287   0.0034   0.0001   0.1402   0.0001
              7049     6501    11066     4712    11241

SPHR01     0.01041 -0.01704  0.05502 -0.01619 -0.05406
            0.3330   0.1242   0.0001   0.2547   0.0001
              8648     7971    13514     4952    13772

SPHR03     0.00426 -0.00608  0.04283 -0.01963 -0.05087
            0.6921   0.5873   0.0001   0.1672   0.0001
              8648     7971    13514     4952    13772

SPHR04     0.01033 -0.01705  0.05492 -0.01619 -0.05410
            0.3370   0.1281   0.0001   0.2547   0.0001
              8648     7971    13514     4952    13772

S4HR01     0.00336 -0.00910  0.00947  0.08192 -0.04921
            0.8188   0.5513   0.4159   0.0001   0.0001
              4651     4286     7386     3280     7524

S4HR03    -0.00056 -0.01595  0.01474  0.03892 -0.00405
            0.9697   0.2965   0.2051   0.0258   0.7251
              4651     4286     7386     3280     7524

S4HR04     0.00336 -0.00910  0.00947  0.08192 -0.04921
            0.8188   0.5513   0.4159   0.0001   0.0001
              4651     4286     7386     3280     7524

S2HR01     0.02010  0.00188  0.00253 -0.05219 -0.06991
            0.0985   0.8834   0.7941   0.0004   0.0001
              6755     6238    10610     4581    10788

S2HR03     0.01925  0.00091 -0.02042 -0.01588 -0.10403
            0.1136   0.9426   0.0354   0.2825   0.0001
              6756     6238    10610     4581    10788

S2HR04     0.00680  0.00613  0.02006 -0.08337  0.01400
            0.5762   0.6283   0.0388   0.0001   0.1459
              6756     6238    10610     4581    10788

U2HR01    -0.03157 -0.01105  0.06004 -0.14981  0.11760
            0.0210   0.4406   0.0001   0.0001   0.0001
              5343     4875     8405     3715     8555

U2HR03    -0.01866 -0.01507  0.05972 -0.14622  0.12686
            0.1726   0.2927   0.0001   0.0001   0.0001
              5343     4475     8405     3715     8555

U2HR04    -0.03069 -0.00707  0.05867 -0.15632  0.12957
            0.0244   0.6218   0.0001   0.0001   0.0001
              5343     4875     8405     3715     8555
```

Reproduced from best available copy.

Table 3-6. (continued)

```
                  CORRELATION COEFFICIENTS / PROB > |R| UNDER HO:RHO=0 / NUMBER OF OBSERVATIONS

                ALDE  INDOCCAV  CROWDING  POLLUTNF    PCDOCB

CONRO1      -0.03138  -0.00662   0.00375  -0.12838   0.22128
              0.0092    0.5983    0.6968    0.0001    0.0001
                6880      6339     10786      4503     10962

CONRO3      -0.02541   0.00425  -0.00027  -0.12647   0.19589
              0.0351    0.7353    0.9777    0.0001    0.0001
                6880      6339     10786      4503     10962

CONRO4      -0.03205  -0.01720   0.01396  -0.13760   0.24372
              0.0078    0.1709    0.1471    0.0001    0.0001
                6880      6339     10786      4503     10962

AVMAXTMP    -0.00523   0.00610   0.04452   0.20438   0.00304
              0.6222    0.5814    0.0001    0.0001    0.7173
                8873      8167     13890      5014     14155

AVMINTMP    -0.00898   0.00789   0.04865   0.20442   0.06666
              0.3974    0.4759    0.0001    0.0001    0.0001
                8873      8167     13890      5014     14155

DMAXTEMP     0.00617   0.00108   0.03177   0.22360   0.00494
              0.5609    0.9219    0.0002    0.0001    0.5648
                8873      8167     13890      5014     14155

DMINTEMP    -0.01327   0.00423  -0.00880  -0.16761   0.02575
              0.2112    0.7025    0.2998    0.0001    0.0022
                8873      8167     13890      5014     14155

HUMIDRF      0.02005   0.00259  -0.01683   0.05522  -0.01656
              0.0595    0.8154    0.0478    0.0001    0.0494
                8833      8132     13823      4983     14088

AVPRECIP     0.00545   0.00018  -0.01186   0.20808   0.00316
              0.6081    0.9870    0.1623    0.0001    0.7074
                8873      8167     13890      5014     14155

EDCOMCUM    -0.43084   0.32669  -0.14671  -0.00528   0.05655
              0.0001    0.0001    0.0001    0.7127    0.0001
                9015      8297     13966      4929     14204

INCOMCUM    -0.18615   0.12508   0.00967  -0.00095   0.02330
              0.0001    0.0001    0.2747    0.9495    0.0080
                8335      7684     12763      4460     12963

RACEWHO     -0.10162  -0.03721  -0.12859  -0.04812  -0.02885
              0.0001    0.0008    0.0001    0.0007    0.0006
                8896      8190     13888      4905     14155

SEXMFU       0.22899  -0.08156   0.01707   0.00652   0.00264
              0.0001    0.0001    0.0421    0.6444    0.7511
                9066      8345     14171      5014     14441
```

Table 3-6. (continued)

```
                  CORRELATION COEFFICIENTS / PROB > |R| UNDER HO:RHO=0 / NUMBER OF OBSERVATIONS

                 BLUE  INDOCCAV  CROWDING  POLLEINF    PCOUCH

AGE          -0.06926   0.11251  -0.32765  -0.00033   0.00232
               0.0001    0.0001    0.0001    0.9816    0.7801
                 9066      8345     14171      5014     14441

AGESQ        -0.05828   0.09217  -0.32086   0.00066   0.00082
               0.0001    0.0001    0.0001    0.9626    0.9214
                 9066      8345     14171      5014     14441

FAT           0.14127  -0.01816  -0.00308   0.02566  -0.03126
               0.0001    0.0997    0.7168    0.0729    0.0002
                 8937      8217     13854      4892     14095

FATSQ         0.13269  -0.01778   0.00153   0.02762  -0.03073
               0.0001    0.1070    0.8572    0.0534    0.0003
                 8937      8217     13854      4892     14095

MANYIND      -0.01335   0.06132   0.04870   0.01027  -0.03591
               0.2038    0.0001    0.0001    0.4672    0.0001
                 9066      8345     14171      5014     14441

VETYIND       0.04365   0.05375  -0.05664   0.00465  -0.01027
               0.0001    0.0001    0.0001    0.7494    0.2214
                 8969      8253     13933      4924     14174

SMOKYIND      0.08951  -0.08224   0.04113   0.01323  -0.00747
               0.0001    0.0001    0.0001    0.3631    0.3844
                 8431      7768     13313      4726     13547

NCIGSDYR      0.07004  -0.06018   0.01801   0.00718  -0.01080
               0.0001    0.0001    0.0400    0.6256    0.2143
                 8238      7593     13003      4626     13231

CHHLMOUR      0.00574   0.01078  -0.09239   0.00035  -0.01544
               0.5847    0.3248    0.0001    0.9805    0.0636
                 9066      8345     14171      5014     14441

CITY1100      0.00903   0.02283   0.02672   0.01368   0.04954
               0.3897    0.0370    0.0015    0.3327    0.0001
                 9066      8345     14171      5014     14441

PSIYIND      -0.01183   0.01044   0.00546  -0.01681   0.15820
               0.3074    0.3872    0.5557    0.2369    0.0001
                 7446      6863     11649      4954     11851

BLUE          1.00000  -0.44016   0.19593   0.00007  -0.03787
               0.0000    0.0001    0.0001    0.9968    0.0003
                 9066      8156      8940      3110      9066

INDOCCAV     -0.44016   1.00000  -0.10407  -0.00684   0.05154
               0.0001    0.0000    0.0001    0.7140    0.0001
                 8156      8345      8230      2877      8348
```

Table 3-6. (continued)

```
                     CORRELATION COEFFICIENTS / PROB > |R| UNDER HO:RHO=0 / NUMBER OF OBSERVATIONS

              PLUE  INDOCCAV CROWDING POLLENHF   PCOOCH

CROWDING   0.19503 -0.10407  1.00000 -0.00954  0.00417
            0.0001   0.0001   0.0000   0.5028   0.6194
              8940     8230    14171     4934    14171

POLLENHF   0.00007 -0.00684 -0.00954  1.00000 -0.15274
            0.9948   0.7140   0.5028   0.0000   0.0001
              7110     2877     4934     5014     5014

PCOOCH    -0.03787  0.05154  0.00417 -0.15274  1.00000
            0.0003   0.0001   0.6194   0.0001   0.0000
              9066     8345    14171     5014    14441
```

Table 3-7. Correlation Coefficients for Sample of Children (N = 15,711)

CORRELATION COEFFICIENTS / PROB > |R| UNDER HO:RHO=0 / NUMBER OF OBSERVATIONS

	O3HR01	O3HR03	O3HR04	SPHR01	SPHR03	SPHR04	S4HR01	S4HR03	S4HR04	S2HR01	S2HR03	S2HR04	N2HR01
O3HR01	1.00000 0.0000 12063	0.92522 0.0001 12063	0.90114 0.0001 12063	0.30954 0.0001 12029	0.22107 0.0001 12029	0.30962 0.0001 12029	0.10808 0.0001 6911	0.10619 0.0001 6911	0.10808 0.0001 6911	-0.21628 0.0001 10362	-0.19510 0.0001 10362	-0.24171 0.0001 10362	0.27990 0.0001 8612
O3HR03	0.92522 0.0001 12063	1.00000 0.0000 12063	0.79499 0.0001 12063	0.29741 0.0001 12029	0.21381 0.0001 12029	0.29755 0.0001 12029	0.11429 0.0001 6911	0.13817 0.0001 6911	0.11429 0.0001 6911	-0.21306 0.0001 10362	-0.19317 0.0001 10362	-0.23206 0.0001 10362	0.31992 0.0001 8612
O3HR04	0.90114 0.0001 12063	0.79499 0.0001 12063	1.00000 0.0000 12063	0.20376 0.0001 12029	0.13353 0.0001 12029	0.20382 0.0001 12029	0.09088 0.0001 6911	0.06154 0.0001 6911	0.09088 0.0001 6911	-0.22064 0.0001 10362	-0.17514 0.0001 10362	-0.24171 0.0001 10362	0.13172 0.0001 8612
SPHR01	0.30954 0.0001 12029	0.29741 0.0001 12029	0.20376 0.0001 12029	1.00000 0.0000 14986	0.91017 0.0001 14986	0.99996 0.0001 14986	0.16924 0.0001 7996	0.16298 0.0001 7996	0.16924 0.0001 7996	0.02499 0.0072 11576	-0.00222 0.8115 11576	-0.01269 0.1720 11576	0.37010 0.0001 8986
SPHR03	0.22107 0.0001 12029	0.21381 0.0001 12029	0.13353 0.0001 12029	0.91017 0.0001 14986	1.00000 0.0000 14986	0.91010 0.0001 14986	0.14299 0.0001 7996	0.17399 0.0001 7996	0.14299 0.0001 7996	0.05629 0.0001 11576	0.03731 0.0001 11576	0.02180 0.0186 11576	0.30070 0.0001 8986
SPHR04	0.30962 0.0001 12029	0.29755 0.0001 12029	0.20382 0.0001 12029	0.99996 0.0001 14986	0.91016 0.0001 14986	1.00000 0.0000 14986	0.16921 0.0001 7996	0.16299 0.0001 7996	0.16921 0.0001 7996	0.02496 0.0072 11576	-0.00226 0.8081 11576	-0.01272 0.1713 11576	0.37045 0.0001 8986
S4HR01	0.10808 0.0001 6911	0.11429 0.0001 6911	0.09088 0.0001 6911	0.16924 0.0001 7996	0.14299 0.0001 7996	0.16921 0.0001 7996	1.00000 0.0000 8007	0.90739 0.0001 8007	1.00000 0.0000 8007	0.26371 0.0001 6689	0.28702 0.0001 6689	0.27737 0.0001 6689	0.01801 0.1834 5462
S4HR03	0.10619 0.0001 6911	0.13817 0.0001 6911	0.06154 0.0001 6911	0.16293 0.0001 7996	0.17399 0.0001 7996	0.16299 0.0001 7996	0.90739 0.0001 8007	1.00000 0.0000 8007	0.90739 0.0001 8007	0.19431 0.0001 6689	0.15150 0.0001 6689	0.21717 0.0001 6689	0.14263 0.0001 5462
S4HR04	0.10808 0.0001 6911	0.11429 0.0001 6911	0.09088 0.0001 6911	0.16924 0.0001 7996	0.14299 0.0001 7996	0.16921 0.0001 7996	1.00000 0.0000 8007	0.90739 0.0001 8007	1.00000 0.0000 8007	0.26371 0.0001 6689	0.28702 0.0001 6689	0.27737 0.0001 6689	0.01801 0.1834 5462

CORRELATION COEFFICIENTS / PROB > |R| UNDER HO:RHO=0 / NUMBER OF OBSERVATIONS

	O3MR01	O3MR03	O3MR04	SPMR01	SPMR03	SPMR04	S4MR01	S4MR03	S4MR04	S2MR01	S2MR03	S2MR04	N2MR01
S2MR01	-0.21628 0.0001 10362	-0.21306 0.0001 10362	-0.22099 0.0001 10362	0.02499 0.0072 11576	0.05829 0.0001 11576	0.02496 0.0072 11576	0.26371 0.0001 6689	0.19931 0.0001 6689	0.26371 0.0001 6689	1.00000 0.0000 11585	0.83587 0.0001 11585	0.86636 0.0001 11585	-0.04938 0.0001 8133
S2MR03	-0.19518 0.0001 10362	-0.19317 0.0001 10362	-0.17518 0.0001 10362	-0.00222 0.8115 11576	0.03731 0.0001 11576	-0.00226 0.8081 11576	0.20702 0.0001 6689	0.15158 0.0001 6689	0.20702 0.0001 6689	0.83587 0.0001 11585	1.00000 0.0000 11585	0.58802 0.0001 11585	-0.09374 0.0001 8133
S2MR04	-0.24171 0.0001 10362	-0.23206 0.0001 10362	-0.24171 0.0001 10362	-0.01269 0.1720 11576	0.02188 0.0186 11576	-0.01272 0.1713 11576	0.27737 0.0001 6689	0.21717 0.0001 6689	0.27737 0.0001 6689	0.86636 0.0001 11585	0.58802 0.0001 11585	1.00000 0.0000 11585	-0.03807 0.0006 8133
N2MR01	0.27990 0.0001 8612	0.31992 0.0001 8612	0.13772 0.0001 8612	0.37018 0.0001 8986	0.36878 0.0001 8986	0.37085 0.0001 8986	0.01801 0.1838 5462	0.14263 0.0001 5462	0.01801 0.1838 5462	-0.04938 0.0001 8133	-0.09374 0.0001 8133	-0.03807 0.0006 8133	1.00000 0.0000 8996
N2MR03	0.22681 0.0001 8612	0.28413 0.0001 8612	0.09973 0.0001 8612	0.32498 0.0001 8986	0.32955 0.0001 8986	0.32517 0.0001 8986	0.01735 0.1997 5462	0.15059 0.0001 5462	0.01735 0.1997 5462	-0.03948 0.0004 8133	-0.03684 0.0005 8133	-0.04334 0.0001 8133	0.90459 0.0001 8996
N2MR04	0.19182 0.0001 8612	0.29196 0.0001 8612	0.04827 0.0001 8612	0.32529 0.0001 8986	0.33131 0.0001 8986	0.32541 0.0001 8986	0.03572 0.0083 5462	0.15426 0.0001 5462	0.03572 0.0083 5462	-0.02070 0.0619 8133	-0.08399 0.0001 8133	0.01998 0.0715 8133	0.95679 0.0001 8996
COMR01	0.01788 0.0637 10755	0.02090 0.0302 10755	-0.04181 0.0001 10755	0.19883 0.0001 11723	0.21219 0.0001 11723	0.19905 0.0001 11723	-0.07946 0.0001 6977	0.00099 0.9342 6977	-0.07946 0.0001 6977	-0.02679 0.0068 10216	-0.03915 0.0001 10216	-0.03693 0.0002 10216	0.39406 0.0001 8743
COMR03	-0.03365 0.0005 10755	-0.02587 0.0073 10755	-0.12161 0.0001 10755	0.15561 0.0001 11723	0.18205 0.0001 11723	0.15581 0.0001 11723	-0.07492 0.0001 6977	-0.00338 0.7777 6977	-0.07492 0.0001 6977	0.00218 0.8260 10216	0.00074 0.9406 10216	-0.02246 0.0232 10216	0.30149 0.0001 8743
COMR04	0.08151 0.0001 10755	0.08857 0.0001 10755	-0.02621 0.0066 10755	0.20100 0.0001 11723	0.20376 0.0001 11723	0.20133 0.0001 11723	-0.04602 0.0001 6977	0.03878 0.0012 6977	-0.04602 0.0001 6977	-0.04010 0.0001 10216	-0.06553 0.0001 10216	-0.02770 0.0050 10216	0.43069 0.0001 8743
AVMAXTMP	0.53700 0.0001 12063	0.53608 0.0001 12063	0.49901 0.0001 12063	0.21821 0.0001 14893	0.14669 0.0001 14893	0.21828 0.0001 14893	0.00298 0.7926 7998	0.01968 0.0791 7998	0.00298 0.7926 7998	-0.37025 0.0001 11585	-0.26295 0.0001 11585	-0.44066 0.0001 11585	0.06902 0.0001 8996
AVMINTMP	0.59006 0.0001 12063	0.55224 0.0001 12063	0.44654 0.0001 12063	0.16189 0.0001 14893	0.09322 0.0001 14893	0.16397 0.0001 14893	0.02803 0.0122 7998	0.04999 0.0001 7998	0.02803 0.0122 7998	-0.38060 0.0001 11585	-0.27808 0.0001 11585	-0.42528 0.0001 11585	0.09912 0.0001 8996
DMAXTEMP	0.50980 0.0001 12063	0.51769 0.0001 12063	0.44997 0.0001 12063	0.14391 0.0001 14893	0.07000 0.0001 14893	0.14393 0.0001 14893	0.17987 0.0001 7998	0.16355 0.0001 7998	0.17987 0.0001 7998	-0.20685 0.0001 11585	-0.14266 0.0001 11585	-0.28287 0.0001 11585	-0.05147 0.0001 8996
DMINTEMP	-0.32069 0.0001 12063	-0.32819 0.0001 12063	-0.32199 0.0001 12063	-0.05412 0.0001 14893	-0.00362 0.6585 14893	-0.05409 0.0001 14893	-0.25020 0.0001 7998	-0.20806 0.0001 7998	-0.25020 0.0001 7998	-0.00975 0.2939 11585	-0.01669 0.0690 11585	-0.00313 0.7359 11585	0.13750 0.0001 8996

Reproduced from best available copy.

```
                              CORRELATION COEFFICIENTS / PROB > |R| UNDER HO:RHO=0 / NUMBER OF OBSERVATIONS

             O3HR01    O3HR03    O3HR08    SPHR01    SPHR03    SPHR08    S4HR01    S4HR03    S4HR08    S2HR01    S2HR03    S2HR08    N2HR01

HUMIDNF    -0.14601  -0.12252  -0.14291  -0.21927  -0.18880  -0.21927   0.15302   0.13742   0.15302   0.11021   0.09478   0.14833   0.01800
             0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0883
              11964     11964     11964     14838     14838     14838      7943      7943      7943     11498     11498     11498      8913

AVPRECIP   -0.19128  -0.16680  -0.13015  -0.22781  -0.18552  -0.22785   0.11131   0.11547   0.11131  -0.07478  -0.02628  -0.05259  -0.24444
             0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0047    0.0001    0.0001
              12063     12063     12063     14893     14893     14893      7994      7994      7994     11585     11585     11585      8996

INCOMCON   -0.01832  -0.03044  -0.00920  -0.08256  -0.07220  -0.08265  -0.03972  -0.01554  -0.03972  -0.00953  -0.00546  -0.01510  -0.01801
             0.0563    0.0015    0.3376    0.0001    0.0001    0.0001    0.0007    0.1858    0.0007    0.3311    0.5776    0.1235    0.1054
              10862     10862     10862     13458     13458     13458      7251      7251      7251     10399     10399     10399      8087

RACEWIHO    0.06563   0.09561   0.07602  -0.03590  -0.03983  -0.03576  -0.06490  -0.04057  -0.06490  -0.04001  -0.03372  -0.04476   0.01415
             0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0003    0.0001    0.0001    0.0003    0.0001    0.1852
              11757     11757     11757     14614     14614     14614      7845      7845      7845     11354     11354     11354      8771

SEXWIFO    -0.01832  -0.01945  -0.01871   0.01005   0.01372   0.01009  -0.00600  -0.00522  -0.00600   0.00785   0.00712   0.00759  -0.00149
             0.0442    0.0326    0.0398    0.2191    0.0930    0.2175    0.5917    0.6408    0.5917    0.3983    0.4435    0.4143    0.8875
              12063     12063     12063     14946     14946     14946      8007      8007      8007     11585     11585     11585      8996

AGE         0.01015   0.00851   0.01519  -0.01984  -0.01640  -0.01992   0.02897   0.03013   0.02897   0.01263   0.00516   0.00613  -0.00384
             0.2650    0.3499    0.0953    0.0153    0.0450    0.0149    0.0095    0.0070    0.0095    0.1739    0.5787    0.5098    0.7155
              12063     12063     12063     14946     14946     14946      8007      8007      8007     11585     11585     11585      8996

AGE30       0.01310   0.01182   0.01876  -0.01709  -0.01402  -0.01718   0.02354   0.02558   0.02354   0.01291   0.00579   0.00627  -0.00316
             0.1503    0.1944    0.0393    0.0367    0.0865    0.0358    0.0352    0.0221    0.0352    0.1648    0.5332    0.4999    0.7647
              12063     12063     12063     14946     14946     14946      8007      8007      8007     11585     11585     11585      8996

CHRCONDN   -0.00650  -0.00293  -0.00502  -0.01734  -0.01707  -0.01741  -0.02754  -0.01407  -0.02754  -0.01736  -0.01556  -0.01231   0.00401
             0.4727    0.7472    0.5811    0.0340    0.0364    0.0333    0.0137    0.2082    0.0137    0.0617    0.0936    0.1851    0.7037
              12063     12063     12063     14946     14946     14946      8007      8007      8007     11585     11585     11585      8996

CITYTIOO   -0.07583  -0.04759  -0.04647   0.08269   0.08242   0.08258   0.03155  -0.00748   0.03155   0.04272   0.01075   0.04616   0.08555
             0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0048    0.5034    0.0048    0.0001    0.2474    0.0001    0.0001
              12063     12063     12063     14946     14946     14946      8007      8007      8007     11585     11585     11585      8996

CROWDING    0.08877   0.10170   0.07000   0.10035   0.08655   0.10035   0.05247   0.05075   0.05247  -0.00926  -0.00800  -0.00585   0.04211
             0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.3209    0.3912    0.5309    0.0001
              11978     11978     11978     14828     14828     14828      7959      7959      7959     11496     11496     11496      8924

POLLENRF    0.01222   0.00972   0.02759  -0.00934  -0.01312  -0.00930   0.05628   0.01993   0.05628  -0.04672  -0.02269  -0.07655  -0.15863
             0.3811    0.4859    0.0479    0.4945    0.3353    0.4945    0.0008    0.2378    0.0008    0.0010    0.1095    0.0001    0.0001
               5142      5142      5142      5395      5395      5395      3512      3512      3512      4977      4977      4977      3933

PCOHC       0.01028   0.05393  -0.06221  -0.06026  -0.05711  -0.06034  -0.04467  -0.00215  -0.04467  -0.03893  -0.07713   0.03542   0.13883
             0.2588    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.8476    0.0001    0.0001    0.0001    0.0001    0.0001
              12063     12063     12063     14946     14946     14946      8007      8007      8007     11585     11585     11585      8996
-----------------------------------------------------------------------------------------------------------------------------------------
             N2HR03    N2HR08    COHR01    COHR03    COHR08  AVMAXTMP  AVMINTMP  DMAXITMP  DMINITMP   HUMIDNF  AVPRECIP  INCOMCON  RACEWIHO

O3HR01      0.22661   0.19182   0.01788  -0.03365   0.08151   0.53744   0.51806   0.50984  -0.32069  -0.14601  -0.19128  -0.01832   0.06563
             0.0001    0.0001    0.0637    0.0005    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0563    0.0001
               8612      8612     10755     10755     10755     12063     12063     12063     12063     11964     12063     10862     11757
```

```
                         CORRELATION COEFFICIENTS / PROB > |R| UNDER HO:RHO=0 / NUMBER OF OBSERVATIONS

            N2HR03    N2HR04    COHR01    COHR03    COHR04  AVMAXTMP  AVMINTMP  DMAXTEMP  DMINTEMP   HUMIDRF  AVPRECIP  INCOMLUN  RACEWIND

O3HR03     0.28813   0.24186   0.02090  -0.02587   0.08857   0.53648   0.55224   0.51769  -0.32814  -0.12252  -0.16680  -0.03044   0.04561
            0.0001    0.0001    0.0302    0.0073    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0015    0.0001
              8612      8612     10755     10755     10755     12063     12063     12063     12063     11964     12063     10862     11757

O3HR04     0.09973   0.09827  -0.08181  -0.12161  -0.02621   0.49941   0.48654   0.48997  -0.32194  -0.14291  -0.13015  -0.00920   0.07602
            0.0001    0.0001    0.0001    0.0001    0.0066    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.3376    0.0001
              8612      8612     10755     10755     10755     12063     12063     12063     12063     11964     12063     10862     11757

SPHR01     0.32898   0.32520   0.19883   0.15561   0.20108   0.21821   0.16389   0.14391  -0.05412  -0.21927  -0.22781  -0.08256  -0.03544
            0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001
              8986      8986     11723     11723     11723     14893     14893     14893     14893     14838     14893     13458     14014

SPHR03     0.32955   0.33131   0.21210   0.18205   0.20376   0.18664   0.09322   0.07000  -0.00362  -0.18880  -0.18552  -0.07220  -0.03983
            0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.6585    0.0001    0.0001    0.0001    0.0001
              8986      8986     11723     11723     11723     14893     14893     14893     14893     14838     14893     13458     14014

SPHR04     0.32517   0.32541   0.19905   0.15581   0.20133   0.21628   0.16397   0.14393  -0.05409  -0.21927  -0.22785  -0.08265  -0.03576
            0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001
              8986      8986     11723     11723     11723     14893     14893     14893     14893     14838     14893     13458     14014

S4HR01     0.01735   0.03572  -0.07946  -0.07492  -0.04602   0.00294   0.02803   0.17987  -0.25020   0.15302   0.11131  -0.03972  -0.06490
            0.1997    0.0063    0.0001    0.0001    0.0001    0.7926    0.0122    0.0001    0.0001    0.0001    0.0001    0.0007    0.0001
              5862      5862      6977      6977      6977      7994      7994      7994      7994      7943      7994      7251      7045

S4HR03     0.15059   0.15426   0.00099  -0.00339   0.03878   0.01964   0.04999   0.16355  -0.20806   0.13742   0.11547  -0.01554  -0.04057
            0.0001    0.0001    0.9342    0.7779    0.0012    0.0791    0.0001    0.0001    0.0001    0.0001    0.0001    0.1856    0.0003
              5862      5862      6977      6977      6977      7994      7994      7994      7994      7943      7994      7251      7045

S4HR04     0.01735   0.03572  -0.07946  -0.07492  -0.04602   0.00294   0.02803   0.17987  -0.25020   0.15302   0.11131  -0.03972  -0.06490
            0.1997    0.0063    0.0001    0.0001    0.0001    0.7926    0.0122    0.0001    0.0001    0.0001    0.0001    0.0007    0.0001
              5862      5862      6977      6977      6977      7994      7994      7994      7994      7943      7994      7251      7045

S2HR01    -0.03948  -0.02070  -0.02679   0.00214  -0.04010  -0.37025  -0.38060  -0.20685  -0.00975   0.11621  -0.07478  -0.00953  -0.04001
            0.0004    0.0619    0.0068    0.8260    0.0001    0.0001    0.0001    0.0001    0.2939    0.0001    0.0001    0.3311    0.0001
              8133      8133     10216     10216     10216     11585     11585     11585     11585     11498     11585     10399     11354

S2HR03    -0.03889  -0.08399  -0.03915   0.00074  -0.06553  -0.26295  -0.27808  -0.14268  -0.01689   0.09478  -0.02628  -0.00546  -0.03372
            0.0005    0.0001    0.0001    0.9406    0.0001    0.0001    0.0001    0.0001    0.0690    0.0001    0.0047    0.5778    0.0003
              8133      8133     10216     10215     10216     11585     11585     11585     11585     11498     11585     10399     11354

S2HR04    -0.08330   0.01998  -0.03693  -0.02246  -0.02776  -0.40066  -0.42528  -0.24247  -0.00313   0.14433  -0.05239  -0.01510  -0.04476
            0.0001    0.0715    0.0002    0.0232    0.0050    0.0001    0.0001    0.0001    0.7354    0.0001    0.0001    0.1235    0.0001
              8133      8133     10216     10216     10216     11585     11585     11585     11585     11498     11585     10399     11354

N2HR01     0.90859   0.95679   0.39406   0.30144   0.43063   0.06702   0.09912  -0.05147   0.13754   0.01806  -0.24848  -0.01801   0.01415
            0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0883    0.0001    0.1054    0.1852
              8996      8996      8743      8743      8743      8996      8996      8996      8996      8913      8996      8087      6771

N2HR03     1.00000   0.83588   0.39970   0.33750   0.42213   0.02882   0.06201  -0.08755   0.16049   0.03460  -0.19214  -0.02455  -0.00804
            0.0000    0.0001    0.0001    0.0001    0.0001    0.0063    0.0001    0.0001    0.0001    0.0011    0.0001    0.0079    0.4513
              8996      8996      8743      8743      8743      8996      8996      8996      8996      8913      8996      8087      6771
```

```
                    CORRELATION COEFFICIENTS / PROB > |R| UNDER HO:RHO=0 / NUMBER OF OBSERVATIONS

           N2NR03    N2NR04    CONR01    CONR03    CONR04  AVMAXTMP  AVMINTMP  DMAXTEMP  DMINTEMP   HUMIDHF  AVPRECIP  INCOMCON  RACEWIHO

N2NR04    0.83548   1.00000   0.37020   0.27565   0.82679  -0.02552   0.01910  -0.11119   0.15940   0.07586  -0.22173  -0.02831  -0.00620
           0.0001    0.0000    0.0001    0.0001    0.0001    0.0155    0.0700    0.0001    0.0001    0.0001    0.0001    0.0109    0.5614
             8996      8996      8743      8743      8743      8996      8996      8996      8996      8913      8996      8087      8771

CONR01    0.39090   0.37020   1.00000   0.92839   0.95169  -0.05968  -0.05138  -0.12947   0.13831  -0.10931  -0.07408   0.00738   0.06347
           0.0001    0.0001    0.0000    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.4485    0.0001
             8743      8743     11745     11745     11745     11692     11692     11692     11692     11646     11692     10560     11451

CONR03    0.33759   0.27565   0.92839   1.00000   0.83738  -0.11437  -0.11035  -0.18269   0.16753  -0.08354  -0.04386  -0.00884   0.04704
           0.0001    0.0001    0.0001    0.0000    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.3856    0.0001
             8743      8743     11745     11745     11745     11692     11692     11692     11692     11646     11692     10560     11451

CONR04    0.82213   0.82679   0.95169   0.83738   1.00000  -0.04947  -0.02342  -0.09622   0.10956  -0.06246  -0.08651  -0.00849   0.06290
           0.0001    0.0001    0.0001    0.0001    0.0000    0.0001    0.0113    0.0001    0.0001    0.0001    0.0001    0.3828    0.0001
             8743      8743     11745     11745     11745     11692     11692     11692     11692     11646     11692     10560     11451

AVMAXTMP  0.02882  -0.02552  -0.05968  -0.11437  -0.04947   1.00000   0.96814   0.72716  -0.24668  -0.36535   0.02046  -0.02470  -0.03632
           0.0063    0.0155    0.0001    0.0001    0.0001    0.0000    0.0001    0.0001    0.0001    0.0001    0.0113    0.0037    0.0001
             8996      8996     11692     11692     11692     15345     15345     15345     15345     15223     15345     13824     15005

AVMINTMP  0.06201   0.01910  -0.05138  -0.11035  -0.02342   0.96814   1.00000   0.72871  -0.24659  -0.23681   0.08370  -0.02279  -0.04697
           0.0001    0.0700    0.0001    0.0001    0.0113    0.0001    0.0000    0.0001    0.0001    0.0001    0.0001    0.0074    0.0001
             8996      8996     11692     11692     11692     15345     15345     15345     15345     15223     15345     13824     15005

DMAXTEMP -0.08755  -0.11119  -0.12947  -0.18269  -0.09622   0.72716   0.72871   1.00000  -0.83652  -0.18212   0.05885   0.00256  -0.01562
           0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0000    0.0001    0.0001    0.0001    0.7615    0.0557
             8996      8996     11692     11692     11692     15345     15345     15345     15345     15223     15345     13824     15005

DMINTEMP  0.16049   0.15940   0.13831   0.16753   0.10956  -0.24668  -0.24659  -0.83652   1.00000   0.02052  -0.04795  -0.01693  -0.00527
           0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0000    0.0114    0.0001    0.0466    0.5187
             8996      8996     11692     11692     11692     15345     15345     15345     15345     15223     15345     13824     15005

HUMIDHF   0.03468   0.07586  -0.10931  -0.08354  -0.06246  -0.36535  -0.23681  -0.18212   0.02052   1.00000   0.20626   0.02435  -0.00871
           0.0011    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0114    0.0000    0.0001    0.0043    0.2869
             8913      8913     11646     11646     11646     15223     15223     15223     15223     15287     15223     13772     14948

AVPRECIP -0.19218  -0.22173  -0.07408  -0.04386  -0.08651   0.02046   0.08370   0.05885  -0.04795   0.20626   1.00000  -0.00752  -0.05645
           0.0001    0.0001    0.0001    0.0001    0.0001    0.0113    0.0001    0.0001    0.0001    0.0001    0.0000    0.3764    0.0001
             8996      8996     11692     11692     11692     15345     15345     15345     15345     15223     15345     13824     15005

INCOMCON -0.02955  -0.02831   0.00738  -0.00884  -0.00849  -0.02470  -0.02279   0.00258  -0.01693   0.02435  -0.00752   1.00000   0.27812
           0.0079    0.0109    0.4485    0.3856    0.3828    0.0037    0.0074    0.7615    0.0466    0.0043    0.3764    0.0000    0.0001
             8087      8087     10560     10560     10560     13824     13824     13824     13824     13772     13824     14165     13862

RACEWIHO -0.00808  -0.00620   0.06347   0.04704   0.06290  -0.03632  -0.04697  -0.01562  -0.00527  -0.00871  -0.05645   0.27812   1.00000
           0.4513    0.5614    0.0001    0.0001    0.0001    0.0001    0.0001    0.0557    0.5187    0.2869    0.0001    0.0001    0.0000
             8771      8771     11451     11451     11451     15005     15005     15005     15005     14948     15005     13862     15362

SEXWIFO   0.00137   0.00157   0.00881   0.01195   0.00009  -0.01190  -0.01568  -0.01450   0.00993  -0.01334   0.00152  -0.00474   0.00807
           0.8969    0.8820    0.3397    0.1955    0.9922    0.1400    0.0521    0.0725    0.2188    0.0779    0.8506    0.5682    0.3173
             8996      8996     11745     11745     11745     15345     15345     15345     15345     15287     15345     14165     15362
```

```
                     CORRELATION COEFFICIENTS / PROB > |R| UNDER HO:RHO=0 / NUMBER OF OBSERVATIONS

             N2HR03    N2HR04    COHR01    COHR03    COHR04  AVMAXTMP  AVMINTMP  DMAXTEMP  DMINTEMP   HUMIDHF  AVPRECIP  INCOMCON  MALEHEAD

AGE        -0.00178  -0.00239   0.01084   0.00507   0.01044   0.00629   0.01159   0.02022  -0.02162  -0.00819   0.00784   0.15282   0.02017
             0.8658    0.8204    0.2401    0.5825    0.2577    0.4360    0.1569    0.0122    0.0074    0.3112    0.3310    0.0001    0.0124
               8996      8996     11745     11745     11745     15345     15345     15345     15345     15267     15345     14165     15362

AGESQ      -0.00113  -0.00247   0.00839   0.00352   0.00777   0.00827   0.01220   0.02269  -0.02473  -0.01043   0.00654   0.14872   0.02352
             0.9148    0.8148    0.3632    0.7030    0.3999    0.3054    0.1309    0.0046    0.0022    0.1973    0.4178    0.0001    0.0036
               8996      8996     11745     11745     11745     15345     15345     15345     15345     15267     15345     14165     15362

CHRLMOUN    0.00597   0.00606   0.00056   0.00547   0.00624   0.01899   0.02228   0.01257  -0.00252   0.00674   0.01955  -0.03451   0.01041
             0.5711    0.5654    0.9518    0.5535    0.4990    0.0187    0.0058    0.1193    0.7547    0.4050    0.0155    0.0001    0.1968
               8996      8996     11745     11745     11745     15345     15345     15345     15345     15267     15345     14165     15362

CITYTIOU    0.03892   0.04811   0.00893   0.02033   0.00199  -0.00498  -0.06519  -0.00990   0.00480  -0.02778  -0.01994  -0.21306  -0.32081
             0.0002    0.0001    0.3331    0.0276    0.8289    0.5371    0.5203    0.2201    0.5524    0.0006    0.0135    0.0001    0.0001
               8996      8996     11745     11745     11745     15345     15345     15345     15345     15267     15345     14165     15362

CROWDING    0.07919   0.08143   0.00416   0.00382   0.02236   0.08131   0.08262   0.03810   0.01087  -0.01417  -0.01264  -0.21642  -0.13998
             0.0001    0.0001    0.6537    0.6803    0.0158    0.0001    0.0001    0.0001    0.1797    0.0809    0.1189    0.0001    0.0001
               8924      8924     11653     11653     11653     15231     15231     15231     15231     15165     15231     14106     15243

POLLENRF   -0.14621  -0.17734  -0.13661  -0.13853  -0.14397   0.20854   0.21060   0.24390  -0.18545   0.06076   0.18062  -0.00644  -0.03708
             0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.0001    0.6514    0.0057
               3933      3933      4889      4869      3889      5468      5468      5468      5468      5426      5468      4922      5332

PCDUC       0.14562   0.15363   0.22698   0.19772   0.24851  -0.02337   0.04309  -0.01282   0.03115  -0.02332   0.00450   0.02295  -0.04510
             0.0001    0.0001    0.0001    0.0001    0.0001    0.0038    0.0001    0.1124    0.0001    0.0039    0.5772    0.0063    0.0001
               8996      8996     11745     11745     11745     15345     15345     15345     15345     15267     15345     14165     15362
--------------------------------------------------------------------------------------------------------------------------------------
            SEXMIFO       AGE     AGESQ  CHRLMOUN  CITYTION  CROWDING  POLLENRF     PCDUC

O3HR01     -0.01832   0.01015   0.01310  -0.00654  -0.07583   0.08877   0.01222   0.01026
             0.0442    0.2650    0.1503    0.4727    0.0001    0.0001    0.3841    0.2580
              12063     12063     12063     12063     12063     11974      5142     12063

O3HR03     -0.01945   0.00851   0.01182  -0.00293  -0.04759   0.10170   0.00972   0.05393
             0.0326    0.3499    0.1944    0.7472    0.0001    0.0001    0.4859    0.0001
              12063     12063     12063     12063     12063     11974      5142     12063

O3HR04     -0.01871   0.01519   0.01876  -0.00502  -0.04647   0.07000   0.02759  -0.06221
             0.0398    0.0953    0.0393    0.5811    0.0001    0.0001    0.0479    0.0001
              12063     12063     12063     12063     12063     11974      5142     12063

SPHR01      0.01005  -0.01989  -0.01709  -0.01731   0.08269   0.10045  -0.00930  -0.06026
             0.2191    0.0153    0.0367    0.0340    0.0001    0.0001    0.4945    0.0001
              14946     14946     14946     14946     14946     14828      5395     14946

SPHR03      0.01372  -0.01680  -0.01402  -0.01707   0.08292   0.06655  -0.01312  -0.05711
             0.0934    0.0450    0.0865    0.0369    0.0001    0.0001    0.3353    0.0001
              14946     14946     14946     14946     14946     14828      5395     14946

SPHR04      0.01009  -0.01992  -0.01718  -0.01741   0.08258   0.10035  -0.00930  -0.06034
             0.2175    0.0149    0.0358    0.0333    0.0001    0.0001    0.4945    0.0001
              14946     14946     14946     14946     14946     14828      5395     14946
```

Reproduced from best available copy.

```
                      CORRELATION COEFFICIENTS / PROB > |R| UNDER HO:RHO=0 / NUMBER OF OBSERVATIONS

            SEXMIFD      AGE     AGESQ CHRLNDUM  CITYIDN  SMOKING POLLEXPF     PCDUC

S4NH01     -0.00600  0.02847  0.02354 -0.02754  0.03155  0.05247  0.05628 -0.04467
             0.5917   0.0095   0.0352   0.0137   0.0048   0.0001   0.0008   0.0001
               8007     8007     8007     8007     8007     7959     3512     8007

S4NH03     -0.00522  0.03013  0.02558 -0.01407 -0.00744  0.05075  0.01991 -0.00215
             0.6408   0.0070   0.0221   0.2082   0.5034   0.0001   0.2378   0.8476
               8007     8007     8007     8007     8007     7959     3512     8007

S4NH04     -0.00600  0.02847  0.02354 -0.02754  0.03155  0.05247  0.05628 -0.04467
             0.5917   0.0095   0.0352   0.0137   0.0048   0.0001   0.0008   0.0001
               8007     8007     8007     8007     8007     7959     3512     8007

S2NH01      0.00765  0.01263  0.01291 -0.01736  0.04272 -0.00926 -0.04672 -0.03093
             0.3983   0.1739   0.1648   0.0617   0.0001   0.3209   0.0010   0.0001
              11585    11585    11585    11585    11585    11490     4977    11585

S2NH03      0.00712  0.00516  0.00579 -0.01558  0.01075 -0.00800 -0.02269 -0.07713
             0.4435   0.5787   0.5332   0.0936   0.2474   0.3912   0.1095   0.0001
              11585    11585    11585    11585    11585    11490     4977    11585

S2NH04      0.00759  0.00613  0.00627 -0.01231  0.08016 -0.00585 -0.07655  0.03542
             0.4143   0.5098   0.4999   0.1851   0.0001   0.5309   0.0001   0.0001
              11585    11585    11585    11585    11585    11490     4977    11585

N2NH01     -0.00149 -0.00389 -0.00316  0.00401  0.04555  0.08211 -0.15883  0.13883
             0.8875   0.7155   0.7647   0.7037   0.0001   0.0001   0.0001   0.0001
               8996     8996     8996     8996     8996     8924     3933     8996

N2NH03      0.00137 -0.00178 -0.00113  0.00597  0.03892  0.07519 -0.14621  0.14562
             0.8967   0.8654   0.9148   0.5711   0.0002   0.0001   0.0001   0.0001
               8996     8996     8996     8996     8996     8924     3933     8996

N2NH04      0.00157 -0.00239 -0.00247  0.00606  0.04811  0.08143 -0.17734  0.15363
             0.8820   0.8204   0.8148   0.5654   0.0001   0.0001   0.0001   0.0001
               8996     8996     8996     8996     8996     8924     3933     8996

CONH01      0.00881  0.01084  0.00839  0.00056  0.00893  0.00416 -0.13661  0.22698
             0.3397   0.2401   0.3632   0.9514   0.3331   0.6537   0.0001   0.0001
              11745    11745    11745    11745    11745    11653     4889    11745

CONH03      0.01195  0.00507  0.00352  0.00547  0.02033  0.00382 -0.13653  0.19772
             0.1955   0.5825   0.7030   0.5535   0.0276   0.6803   0.0001   0.0001
              11745    11745    11745    11745    11745    11653     4889    11745

CONH04      0.00009  0.01094  0.00777  0.00624  0.00199  0.02236 -0.14397  0.24851
             0.9922   0.2577   0.3999   0.4990   0.8289   0.0156   0.0001   0.0001
              11745    11745    11745    11745    11745    11653     4889    11745

AVMAXTMP   -0.01190  0.00629  0.00827  0.01899 -0.00498  0.08131  0.20854 -0.02337
             0.1404   0.4360   0.3059   0.0187   0.5371   0.0001   0.0001   0.0038
              15345    15345    15345    15345    15345    15231     5468    15345
```

```
                    CORRELATION COEFFICIENTS / PROB > |R| UNDER HO:RHO=0 / NUMBER OF OBSERVATIONS

          SEXM1F0       AGE     AGESQ  CHRLMDUM  CITY1100  CROWDING  POLLENRF     PCHUC

AVMINTMP -0.01568   0.01159   0.01220   0.02228  -0.00519   0.08262   0.21060   0.04309
           0.0521    0.1509    0.1309    0.0058    0.5203    0.0001    0.0001    0.0001
            15345     15345     15345     15345     15345     15231      5468     15345

DMAXTEMP -0.01450   0.02022   0.02289   0.01257  -0.00990   0.03810   0.24390  -0.01282
           0.0725    0.0122    0.0046    0.1193    0.2201    0.0001    0.0001    0.1124
            15345     15345     15345     15345     15345     15231      5468     15345

DMINTEMP  0.00993  -0.02162  -0.02473  -0.00252   0.00480   0.01087  -0.18545   0.03115
           0.2189    0.0074    0.0022    0.7587    0.5524    0.1797    0.0001    0.0001
            15345     15345     15345     15345     15345     15231      5468     15345

HUMIDHF  -0.01339  -0.00819  -0.01043   0.00674  -0.02778  -0.01417   0.06076  -0.02332
           0.0979    0.3112    0.1973    0.4050    0.0006    0.0809    0.0001    0.0039
            15287     15287     15287     15287     15287     15165      5426     15287

AVPRECIP  0.00152   0.00784   0.00654   0.01955  -0.01998  -0.01264   0.18062   0.00450
           0.8506    0.3316    0.4178    0.0155    0.0135    0.1189    0.0001    0.5772
            15345     15345     15345     15345     15345     15231      5468     15345

INCOMCON -0.00479   0.15282   0.14872  -0.03951  -0.21368  -0.21642  -0.00644   0.02295
           0.5682    0.0001    0.0001    0.0001    0.0001    0.0001    0.6514    0.0063
            14165     14165     14165     14165     14165     14108      4922     14165

RACEW1B0  0.00807   0.02017   0.02352   0.01041  -0.32681  -0.13998  -0.03788  -0.04510
           0.3173    0.0124    0.0036    0.1968    0.0001    0.0001    0.0057    0.0001
            15362     15362     15362     15362     15362     15243      5332     15362

SEXM1F0   1.00000  -0.00870  -0.00857   0.02252  -0.00690  -0.01004  -0.00559   0.00068
           0.0000    0.2756    0.2827    0.0048    0.3872    0.2100    0.6797    0.9320
            15711     15711     15711     15711     15711     15589      5468     15711

AGE      -0.00870   1.00000   0.96683   0.05740  -0.02571  -0.01695  -0.02870   0.01091
           0.2756    0.0000    0.0001    0.0001    0.0013    0.0343    0.0336    0.1716
            15711     15711     15711     15711     15711     15589      5468     15711

AGESQ    -0.00857   0.96683   1.00000   0.05070  -0.02255  -0.02031  -0.02761   0.00863
           0.2827    0.0001    0.0000    0.0001    0.0047    0.0112    0.0412    0.2792
            15711     15711     15711     15711     15711     15589      5468     15711

CHRLMDUM  0.02252   0.05740   0.05070   1.00000   0.00601  -0.00907  -0.01995   0.00519
           0.0048    0.0001    0.0001    0.0000    0.4513    0.2575    0.1403    0.5153
            15711     15711     15711     15711     15711     15589      5468     15711

CITY1100 -0.00690  -0.02571  -0.02255   0.00601   1.00000   0.08226   0.00118   0.03814
           0.3872    0.0013    0.0047    0.4513    0.0000    0.0001    0.9305    0.0001
            15711     15711     15711     15711     15711     15589      5468     15711

CROWDING -0.01004  -0.01695  -0.02031  -0.00907   0.08226   1.00000  -0.00801  -0.00785
           0.2100    0.0343    0.0112    0.2575    0.0001    0.0000    0.5555    0.3269
            15589     15589     15589     15589     15589     15589      5420     15589
```

Reproduced from
best available copy.

```
                                  CORRELATION COEFFICIENTS / PROB > |R| UNDER HO:RHO=0 / NUMBER OF OBSERVATIONS

          SEXMIFD       AGE     AGESQ CHRLMOUN  CITY100 CROSSING POLLENFF     PCDOC

POLLENFF -0.00559  -0.02870  -0.02761 -0.01995  0.00116 -0.00801  1.00000  -0.15934
           0.6797    0.0338    0.0412   0.1403   0.9304   0.5555   0.0000    0.0001
             5468      5468      5468     5468     5468     5420     5468      5468

PCDOC     0.00068   0.01091   0.00863  0.00519  0.03814 -0.00785 -0.15934   1.00000
           0.9360    0.1716    0.2792   0.5153   0.0001   0.3269   0.0001    0.0000
            15711     15711     15711    15711    15711    15587     5468     15711
```

variables used in our analysis of chronic morbidity is available from the authors.)

For example, consider the simple correlation between the average daily maximum 1-hour ozone concentration for the 14 day reference period (O3NR01) and the average sulfate reading during the same two week period (S4NR01). Table 3-6 indicates that the correlation coefficient is 0.119, that this is significant at better than the 1 percent level (p=.001), and that this coefficient was calculated over the 6,575 individuals for whom ozone and sulfate data were available during their two-week reference period. The corresponding correlation coefficient between ozone and sulfur dioxide (S2NR01) is -0.199, calculated over 9,764 pairs of readings.

It is clear from the length of Tables 3-6 and 3-7 that we cannot discuss the various associations between all the independent variables used in this report. It is important to point out that collinearity does not appear to be a serious problem with respect to the air pollution data insofar as pairwise comparisons are concerned. The highest correlation coefficient between ozone and any other pollutant is 0.281 and it is below 0.20 in absolute value for three of the other pollutants. Ozone is highly correlated with average maximum temperature during the reference period, however, the correlation coefficient being 0.53. This is to be expected since ozone is created when hydrocarbons and nitrogen dioxide interact in the presence of strong sunlight, a condition conducive to high temperatures as well. Other correlation coefficients are available from Tables 3-6 and 3-7.

Because of the great number of different models estimated in this project, we did not print out a separate matrix of correlation coefficients for each and every regression. All models of adult acute morbidity were estimated on subsets of the 14,441 individuals used to calculate the matrix in Table 3-6. Similarly, regressions explaining children's morbidity were based on subsets of the 15,711 children in the overall sample. We have printed out the correlation coefficients for several different subsets of our final sample and the resulting data are very similar to those presented here.

In spite of our assertion that collinearity appears not to be a serious problem with our data, we have performed diagnostic tests on a number of the regressions estimated below using sophisticated tests which permit the detection of multicollinearity when it arises from relationships among more than two regressors. This is not possible from an examination of simple (bivariate) correlation coefficients. A description of these diagnostic techniques, a list of the models to which they were applied, and the results of the tests are presented in Section 4.5. below.

3.3 Estimation Techniques

Given the different measures of health status used as dependent variables in this study, there are many statistical estimation techniques that could be employed. From this number, we have tended in this analysis to employ basic techniques when confronted with several alternatives. The reasons for this are several.

First, as explained below, we have deliberately chosen in this report to explore as many hypotheses as possible in order to make use of our unique data base. This broad approach meant that we could not explore in great detail any one finding, thus our choice of fairly straightforward techniques. Second, the methods utilized in this report tend to be less computationally burdensome than some of the highly sophisticated maximum likelihood estimation (MLE) methods described below. All estimation results presented in this report were computed using the Statistical Analysis System (SAS). While SAS has proven to be quite powerful in terms of handling the large data sets used in this study, it is also somewhat limited in its MLE routines.

We turn now to a discussion of the various estimation techniques considered in our analysis, including those actually utilized and those that seem potentially fruitful for further research within the areas of concern of this study.

The available measures of acute health status--RADs, WLDs, SLDs, and BDDs--raise several questions about estimation methodology. First, data for RADs, etc., as collected in the HIS are discrete, assuming only one of fifteen possible values in $\{0,1,\ldots,14\}$. Days where individuals report more than half of the day as limited due to some illness or injury are coded as one day, with less than or equal to half the day coded as zero. In the present study, we treat the "days" measures of health status as continuous on the interval $[0,14]$, recognizing that there exist econometric techniques to employ in the analysis of such count data. Others have taken some initial steps in this latter direction, utilizing poisson regression in the analysis of the effects of air pollution on work loss days.[5]

Consider now the use of ordinary least squares (OLS) in the analysis of "days" measures of health status. As a brief sketch, consider a dependent variable vector y of dimension nx1, a regressor matrix X of dimension nxk, a parameter vector β of dimension kx1 and stochastic vector of dimension nx1. The econometric model can be written as:

$$y = X\beta + \varepsilon \tag{1}$$

and the OLS estimator of β is

$$\hat{\beta} = (X^TX)^{-1}X^Ty \tag{2}$$

This model is said to satisfy full ideal conditions[5] if

1) X is nonstochastic of rank $k<t$ and $\lim(\bar{X}^TX)/n$ is a finite and nonsingular matrix

and

2) ε is multivariate normal. $\varepsilon \sim N(0. \sigma^2 I_n)$

Given these assumptions (or a slightly weaker set of assumptions where the normality of the ε distribution is not required). we know that the estimator $\hat{\beta}$ of β is unbiased ($E(\hat{\beta}) = \beta$) and consistent ($\operatorname{plim}(\hat{\beta}) = \beta$).

Consider now the application of OLS to estimation of health effects as quantified by RADs. BDDs, WLDs. and SLDs. As mentioned above, we ignore the discrete nature of such data and treat such measures as continuous on [0,14]. What cannot be so easily ignored, however, is the fact that the days measures--the y vector in (1) above--has a range restricted to the interval [0,14]. One is therefore confronted with the existence of a limited dependent variable. That is, for all possible values in the domain X, the y values are constrained to be in [0,14]. One could view the problem perfectly analogously by considering the measure of health status as "percentage of the two-week period restricted, in bed, at home from work, etc.," and view y as restricted to the range [0,1].

Such range restrictions have definite implications for statistical estimation of the parmeter vector β in (1). In particular, for finite X and β, we have a situation where the restriction on the range of y implies a restriction on the set of real (vector) values that can be assumed by ε. Specifically, for all i, ε must be in [$-X_i\beta$, $14-X_i\beta$], i.e., for continuous and differentiable $f(\varepsilon_i)$,

$$\int_{-X_i\beta}^{14-X_i\beta} f(\varepsilon_i)\,d\varepsilon_i = 1,$$

where $f(\varepsilon_i)$ is the density of ε_i.

This restriction implies that certain specifications of the joint density of the ε_i are technically inadmissable, the multivariate normal being the most obvious. Obtaining an unbiased estimator $\hat{\beta}$ of β is still possible with OLS as long as $E(\varepsilon)=0$, but the statistical tests based on the assumption of ε being distributed multivariate normal are no longer strictly appropriate. However, if the full ideal conditions are met, save for the normality assumption, the usual test procedures based on normality can be asymptotically justified so long as $E(\varepsilon)=0$, $Var(\varepsilon)$ is finite, X is uniformly bounded, and lim $(X^TX)/n$ is finite and nonsingular.[7]

Our estimation of the "days" models rest on several assumptions. First, the large number of observations in the models estimated allows us to invoke the asymptotic properties of the estimators. Second, it cannot be ruled out ex ante that $E(\varepsilon)=0$, although given that there exists a considerable mass point at $\varepsilon_i=-X_i\beta$ (i.e. no days reported), actual specification of the density of ε is not attempted here. Finally, and quite practically, computational costs are much lower for OLS estimation as compared with MLE methods, which require the further step of specification of the form of the density of ε. We now concern ourselves with discussion of other methods appropriate to alternative specifications of (1) above. Some of these are in fact utilized in our analysis, others are presented to stimulate thinking about further research beyond the bounds of this study. For now, we merely state that OLS is widely used in the present study; where used, one must be aware of the limitations described above.

An approach considered frequently in econometric analysis of limited dependent variables is the use of latent variables as censoring agents.[8] That is to say, we consider some continuous variable y^* that is meaningfully modeled as

$$y^* = Z\beta + \varepsilon \qquad (3)$$

where ε is well-behaved in the classical sense. For simplicity, we will assume that ε is distributed independent, multivariate normal, although this is an inessential restriction for many results presented. The important thing to note is that there are no restrictions imposed on the domain of ε; we assume only that $f(\varepsilon)$ obeys the usual restrictions on probability density functions.

However, in the limited dependent variable model, we do not observe y^*, or at least we do not observe y^* over its entire range. Rather, we observe some variable y that is related to y^* by the following quite general relationship:

$$\begin{aligned} y &= y^* \text{ if } c_l < y^* < c_u, \\ &= c_l \text{ if } y^* \leq c_l \\ &= c_u \text{ if } y^* \geq c_u, \end{aligned}$$

where $c_l \in [-\infty, c_u)$ and $c_u \in (c_l, +\infty]$ are constants. The problem of the analyst is to estimate the parameter vector β in (3) in a consistent and efficient manner.

It is clear that such a model is germain to estimating the determinants of the "days" measures of individuals' health status. Note that the presentation of the model as a latent variable model differs from our discussion above regarding OLS estimation. Insofar as the "days" measures are concerned, the definition of y^* is not important.

For present purposes, y^* is something like "poorness of health," where, loosely speaking, larger values of y^* imply increasing probability of additional RADs, WLDs, BDDs, or SLDs. The bounds on the observability of y^* are those imposed by the nature of the "days" data, i.e. no fewer than zero and no greater than fourteen "days" can be reported in a two-week period.

If one decides to take a latent variable approach to estimating "days" equations, there are several routes that can be taken. Some analysts have suggested that because of the considerable mass point at zero days, and because there are so few individuals who report upwards of fourteen days of restriction, a tobit model would be an appropriate one to estimate. In the classical tobit model. c_u is irrelevant ($c_u = +\infty$ in terms of the above presentation) and c_l is set at zero. The likelihood function for the tobit model is

$$L = \prod_{S_1} \sigma^{-1} f((y_i - x_i \beta)/\sigma) \prod_{S_2} F(-x_i \beta / \sigma) \qquad (4)$$

where S_1 and S_2 are the noncensored and censored sets of observations. respectively. and $f(\cdot)$ and $F(\cdot)$ are the density and distribution functions of a standard normal variate (we here assume that ε in (3) is distributed independent normally). The parameter vector (β, σ) can be estimated by means of maximum likelihood or by an iterative least squares procedure.[9]

Recognition of the limitation at fourteen days, however, suggests that a two-limit probit estimator be used.[10] Here, c_l and c_u are finite, and for purposes of exposition, it is assumed that there exist observations at c_l and c_u and in-between, as is the case in the "days" data. The likelihood function for this model can be written as

$$L = \prod_{S_1} F((c_l - x_i \beta)/\sigma) \prod_{S_2} F((x_i \beta - c_u)/\sigma) \prod_{S_3} \sigma^{-1} f((y_i - x_i \beta)/\sigma) \qquad (5)$$

S_1. S_2, and S_3 are the sets of observations at the lower and upper bounds, and in-between. respectively. Estimation of β is again performed by maximum likelihood. Again, normality and independence of the error structure in (3) are assumed.

As mentioned earlier. yet another approach would be an explicit recognition of the discrete nature of the data, with estimation under the assumption of a poisson distribution for the days. However. the mean of a poisson-distributed variate is equal to the variance. thus imposing a sometimes unrealistic moment structure on the empirical distribution of the variates. Furthermore, the censoring at fourteen days again implies that the strict poisson model is not appropriate. However, the dearth of observations at fourteen would most likely imply that a poisson model with explicit recognition of the fourteen day limit would yield parameter estimates hardly different than one in which the censoring limit was imposed in the likelihood function. Again, estimation is accomplished by methods of maximum likelihood, or by a weighted least squares approach.

To summarize, the "days" measures of health status present interesting and challenging estimation problems. Whether it is used or not, the latent variable approach has been the methodology most studied, and, therefore, the one on which the most information is available.

However, many interesting questions concerning individuals' health are of a different type. Does an individual have a certain condition or set of conditions? Was an individual restricted in activity to some extent during a particular period? Because measurement of health status is definitionally difficult, many of the most illuminating questions about health are necessarily framed in such binary or discrete terms. That is, attention is confined to the presence or absence of a particular condition or illness. (This is not without its own problems, since two individuals both reporting chronic sinusitis, for instance, might differ substantially in the severity of those conditions.) Fortunately, much attention has been devoted recently to analysis of discrete outcomes (by discrete outcomes, we mean situations where an outcome occurs or it does not--as in the case of an individual who either has a respiratory disease or does not).[11] Here we present the basics of the analysis of discrete outcomes which is used in much of our analysis below. We restrict our attention to binary choice since that is the approach employed in our study.

The basic idea behind the binary outcome model is that the probability of an outcome (an illness, for instance) depends functionally on a vector of variables or attributes characterizing an individual and his environment, broadly defined. Much of the analysis of discrete outcomes has been cast in the framework of individual utility maximization: an individual chooses the outcome which is higher in his preference ordering. For some analysis, however, the notion of deliberate choice is not a particularly appealing one. Although some might contend that all outcomes are to some extent "chosen," it stretches one's normal understanding of the meaning of "choice" to think of outcomes like chronic cardiovascular or respiratory disease as being "chosen." For present purposes, it is much more straightforward to view outcomes from an ex post vantage point and consider the characteristics of an individual and his environment that might be related to the observed (health) outcomes, and how they are related.

In this study binary outcomes have been examined in both the acute and chronic analyses undertaken. For instance, acute health outcomes of the binary nature that have been studied are the presence or absence of restricted-activity, work or school loss, and bed-disability days. Our analysis of chronic morbidity is based on the presence or absence of chronic respiratory, cardiovascular, and several other conditions.

We describe the probability that individual i has outcome y_i (at least one bed disability day, for example) to be a function of a vector of characteristics X_i and a parameter vector β. That is

$$Pr(y_i = \text{illness}) = F(X_i\beta) \tag{5}$$

For convenience, we define $y_i=1$ if the binary outcome is a "yes" or illness and $y_i=0$ if the outcome is "no" or health, and assume that all individuals are characterized by either an illness or health, but not both.

The fact that we are examining probabilities of outcomes implies a special restriction on the range of the function F, i.e., $F(X_i\beta) \in [0,1]$. There is an obvious class of functional forms that definitionally imposes such a restriction on the functional range, however, and this class is the distribution functions corresponding to continuous probability densities. We assume, therefore, that $X_i\beta$ is continuous on $(-\infty, +\infty)$ and examine the nature of the functions $F(X_i\beta)$ such that $F:\mathbf{R}\text{-->}[0,1]$. The main question is: What is the form of F most appropriate for analysis, or, equivalently, what is the appropriate form of the density $f(z)=dF(z)/dz$ of the underlying scalar-valued "indicator" variable $z_i=X_i\beta$?

Although it does not necessarily satisfy (5), the linear probability model has been used frequently in many different area of analysis. This model assumes the form

$$Pr(y_i=1) = X_i\beta, \tag{6}$$

i.e. $F(X_i\ )=X_i\beta$. In order to satisfy (5), (6) is recast as

$$\begin{aligned} Pr(y_i=1) &= 0,\ X_i\beta\ [-\infty, 0] \\ &= X_i\beta,\ X_i\beta\in(0,1) \\ &= 1,\ X_i\beta\in[1,+\infty] \end{aligned} \tag{7}$$

There are several detailed descriptions available of the properties of the linear probability model, the major drawback of which is that predicted probabilities, $X_i \hat{\beta}$, can be outside of $[0,1]$.[12] That is, an individual might be predicted to have greater than a 100 percent chance of having an illness. A corollary problem is that individuals who are predicted to be certain of having a particular condition or illness ($p=1$), may not have that condition at all.

This aside, the linear probability model is straightforward to estimate. One writes

$$y_i = X_i \beta + \varepsilon_i, \tag{8}$$

where y_i is the $n \times 1$ vector of binary (0,1) outcomes. Clearly, ε_i is heteroskedastic. Furthermore, it can assume only one of two possible values, $-X_i \beta$ or $(1-X_i \beta)$. For $E(\varepsilon_i)=0$, one must have $\Pr(\varepsilon_i=-X_i\beta) = (1-X_i \beta)$ and $\Pr(\varepsilon_i=1-X_i \beta)=X_i\beta$. Given this restriction and the fact that the variance of a binomial variate with parameter θ equals $\theta(1-\theta)$, one has $\operatorname{Var}(\varepsilon_i) = X_i \beta(1-X_i \beta)$. Thus, β in (6) or (8) can be estimated by GLS.

$$\tilde{\beta} = (X^T \Omega^{-1} X)^{-1} X^T \Omega^{-1} y \tag{9}$$

where $\Omega = \operatorname{diag}(\operatorname{Var}(\varepsilon_i)) = \operatorname{diag}(X_i \beta(1-X_i \beta))$. Since Ω is unlikely to be known, a feasible GLS (FGLS) estimator would use consistent estimates of $\operatorname{Var}(\varepsilon_i)$ (using, perhaps, $\hat{\beta}_{OLS}$ to predict $X_i\beta$) and have

$$\bar{\bar{\beta}} = (X^T \hat{\Omega}^{-1} X)^{-1} X^T \hat{\Omega}^{-1} y \tag{10}$$

where $\hat{\Omega} = \operatorname{diag}(X_i \hat{\beta}_{OLS}(1-X_i \hat{\beta}_{OLS}))$ for $X_i \hat{\beta}_{OLS} \in (0,1)$. It can be shown that $\operatorname{plim}(X^T \hat{\Omega}^{-1} X)/n$ finite nonsingular and $\operatorname{plim}(X^T \hat{\Omega}^{-1} \varepsilon)/n=0$ are sufficient conditions for the consistency of $\bar{\bar{\beta}}$, and that if $\operatorname{plim}(X^T(\hat{\Omega}^{-1} - \Omega^{-1})X)/n=0$ and $\operatorname{plim}(X^T(\hat{\Omega}^{-1} - \Omega^{-1})\varepsilon)/n^{-.5}=0$ then the FGLS estimator has the same asymptotic distribution as the GLS estimator and, therefore, is efficient.[13] One must be careful when using FGLS, however, as the extremely high weights on observations with predicted probabilities near unity or near zero can cause considerably more variability in parameter estimation than would result from unweighted OLS estimates. This situation is particularly problematic in the small sample case.

The main appeal of the linear probability model rests in its computational simplicity and low cost, in either the OLS or the FGLS cases. We utilized linear probability analysis in the preliminary stages of our analysis, but because of the limitations discussed above, it was not employed in estimation of the models presented in Chapter 4.

The two most popular and most utilized methods for analysis of binary outcomes are the binary probit and binary logit models. In the binary probit model, one has

$$\Pr(y_i=1) = \int_{-\infty}^{X_i\beta/\sigma} f(z)dz = F(X_i\beta/\sigma), \qquad (11)$$

where $z \sim N(0,1)$. The parameter vector β is estimable up to a scalar multiple $(1/\sigma)\beta$ without prior knowledge of σ. In the logit model, one has

$$\Pr(y_i=1) = (1 + \exp(-X_i\beta))^{-1} = F(X_i\beta). \qquad (12)$$

On the assumption that observations are independent draws from a binomial distribution, the likelihood function for these two models, can be written as

$$L(y_i) = \prod_{S_1} F_i^{y_i} \prod_{S_2} (1-F_i)^{(1-y_i)} \qquad (13)$$

where $F_i=F((X_i\beta)/\sigma)$ for binary probit and $F_i=F(X_i\beta)$ for binary logit, and where S_1 and S_2 are the sets of observations of ill and healthly individuals ($y_i=1$ and $y_i=0$), respectively. Estimation of both models is by maximum likelihood, maximizing L in (13) with respect to the parameter vector. Logit modeling has tended to dominate probit in applied analysis because of logit's closed form expression for F. In probit, for each observation at each iteration there is call to the normal distribution function at the value at which the integral in (11) must be evaluated. We have performed the bulk of the binary discrete outcome analysis in the present study in a logit framework, using SAS's PROC LOGIST maximum likelihood routine. In many cases the parameter estimates obtained from MLE of the logit and probit models are quite comparable, up to a multiplicative constant, and there exist formulae for the approximate conversion between logit and probit parameter estimates.[14]

Both the logit and probit models can be estimated by procedures which allow for iterative reweighting of the observations in a nonlinear least squares framework (SAS's PROC NLIN and BMDP's P3R and PAR, for example).[15] Such estimation is by iteratively reweighted nonlinear FGLS (IRNFGLS), based on explicit recognition of the (assumed) independent binomial trial structure of the y_i vector. One has

$$y_i = F(z_i) + \varepsilon_i \tag{14}$$

where the properties of ε_i are basically those discussed above in the case of linear probability models. One iteratively searches for the vector that solves

$$\min\ (y - F)^T \hat{\Omega}^{-1}(y - F) \tag{15}$$

where $\hat{\Omega}^{-1}$ is a diagonal iteratively updated weight matrix with typical element $[(F(X_i\beta_{-1})(1-F(X_i\beta_{-1})))^{-1}]$, where β_{-1} is the previous iteration's estimate of β. It can be shown that IRNFGLS applied to (14) is maximum likelihood given some regularity conditions. Note that since many packages have functions or calls to routines that efficiently calculate cumulative standard normal integrals (e.g. the PROBNORM function in SAS), probit estimation is quite feasible.

To summarize, since many health outcomes of interest are binary in nature, one is confronted with a discrete-outcome model. The above discussion has identified a set of the estimation techniques available for such analysis. Much of the potentially interesting future research might be conducted in the framework of the multi-outcome extensions of the binary outcome models examined here.

Footnotes to Chapter 3

[1] See Atkinson, Crocker, and Murdock [1983].

[2] NCHS [1981], pp. 13-20.

[3] There are two very comprehensive summaries of the effects of ozone and other photochemical oxidant pollution on health. These are U.S. Environmental Protection Agency [1978] and the National Academy of Sciences [1977].

[4] See Goldsmith and Landaw [1968], for instance.

[5] See Hausman, Ostro and Wise [1983].

[6] See Schmidt [1976], p. 2.

[7] See Schmidt [1976], pp. 56-50.

[8] See the surveys by Amemiya [1981] and Maddala [1983].

[9] See Fair [1977].

[10] See Rosett and Nelson [1975].

[11] See the Amemiya [1981] and Maddala [1983] surveys.

[12] See Domenich and McFadden [1975].

[13] See Schmidt [1976].

[14] See Amemiya [1981], pp. 1487-1488.

[15] See Jennrich and Moore [1975].

References to Chapter 3

Amemiya, Takeshi. "Qualitative Response Models: A Survey," *J. Econ. Lit.*, 19(1981), pp. 485-496.

Atkinson, Scott, Thomas Crocker and Robert Murdock, "Have Priors in Aggregate Air Pollution Epidemiology Dictated Posteriors?", working paper from the Department of Economics, University of Wyoming, 1983.

Domencich, Thomas and Daniel McFadden, *Urban Travel Demand* (Amsterdam: North Holland), 1975.

Fair, Ray, "A Note on the Computation of the Tobit Estimator," *Econometrica*, 45(1977), pp. 1723-27.

Goldsmith, John and Stephen Landaw, "Carbon Monoxide and Human Health," *Science*, 162 (1968), pp. 1352-1359.

Hausman, Jerry, Bart Ostro, and David Wise, "Air Pollution and Lost Work," paper presented at the NBER Conference on Productivity in Health, Stanford, California, August 18-19, 1983.

Jennrich, R.I. and R.H. Moore, "Maximum Likelihood Estimation by Means of Nonlinear Least Squares," *American Statistical Association. 1975 Proceedings of the Statistical Computing Section*, pp. 57-65.

Maddala, G.S., *Limited-Dependent and Qualitative Variables in Econometrics* (Cambridge: Cambridge University Press), 1983.

National Academy of Sciences, *Ozone and Other Photochemical Oxidants* (Washington, D.C.: National Academy Press), 1977.

Rosett, R.N. and F.D. Nelson, "Estimation of the Two-Limit Probit Regression Model," *Econometrica*, 43(1975), pp. 141-146.

Schmidt, Peter, *Econometrics* (New York: Marcel Dekker), 1976.

U.S. Environmental Protection Agency, *Air Quality Criteria for Ozone and Other Photochemical Oxidants*, Publication no. EPA-600/8-78-004, April 1978.

Appendix to Chapter 3

DEFINITIONS OF AND PROCEDURES USED TO CREATE MEASURES OF MORBIDITY

The present study, measures of health status, can be separated into two distinct groups. First are the "all-causes" measures of health. Second are more specific measures, dealing primarily with respiratory conditions, but also to a lesser degree with cardiovascular illnesses. Measures in the second group are further disaggregated where appropriate into acute and chronic categories.

In the following descriptions, references are made to definitions presented in Chapter 2 and Appendix A to be as precise as possible about the definitions of the several measures of health status.

Within the all-causes classification, six measures of health status have been utilized. They are:

1) RADMINOR - This is a measure of total two-week restricted-activity days when such restrictions are of a minor nature. This variable is formed as follows:

For adults: RADMINOR = TOTRAD2W if BEDDIS2W = 0 and WKSLLD2W = 0
= n/a if BEDDIS2W > 0 or WKSLLD2W > 0.

For children:

RADMINOR = TOTRAD2W if BEDDIS2W = 0 and WKSLLD2W = 0,
if WKSLLDID = 2
= TOTRAD2W if BEDDIS2W = 0, if WKSLLDID = 1

= n/a if BEDDIS2W > 0 or WKSLLD2W > 0,
if WKSLLDID = 2
= n/a if BEDDIS2W > 0, if WKSLLDID = 1.

The reasons for conditioning the definition of RADMINOR for children on the value of WKSLLDID is, of course, that by definition children of pre-school age will report zero for WKSLLD2W.

2) WKSLLD2W - As defined in Chapter 2, subject to

For adults: WKSLLD2W is defined only for CURACT2W = 1 and
CURACT2W = 2, n/a otherwise.

For children:

WKSLLD2W is defined for WKSLLDID = 2 if PROCQUAR = 1 or PROCQUAR = 4 or PROCQUAR = 2 and WEEK < 10 or PROCQUAR = 3 and WEEK > 9: and n/a otherwise.

For adults, the CURACT2W cutoffs were established so to consider only those individuals for whom work loss is a meaningful concept. The rationale for the cutoffs for children is to proxy as nearly as possible the boundaries of a representative school year. Obviously only those children of school age (WKSLLDID = 2) qualify for analysis in a school-loss days model.

3) BEDDIS2W - As defined in Chapter 2. This variable serves as an aggregate measure of restrictions in activity of a severity sufficient to warrant bed-disability.

4) RADPRESM - This variable is derived from RADMINOR as follows:

RADPRESM = 1 iff RADMINOR > 0

RADPRESM = 0 iff RADMINOR = 0.

This binary measure of minor restricted-activity days is of interest when questions of the "whether or not" or "have, don't have" are posed.

5) BEDPRES - This variable is derived from BEDDIS2W as follows:

BEDPRES = 1 iff BEDDIS2W > 0

BEDPRES = 0 iff BEDDIS2W = 0.

6) WSLPRES - This variable is derived from WKSLLD2W, subject to both the adult and child restrictions described above, as follows:

WSLPRES = 1 iff WKSLLD2W > 0

WSLPRES = 0 iff WKSLLD2W = 0.

Insofar as respiratory conditions are concerned, seven measures have been utilized, six acute and one chronic. The following table illustrates those measures in the Health Interview Survey that we have classified as respiratory conditions. The three-digit condition codes listed are values assumed by the variables DGRCD101, DGRCD102, ... , DGRCD114 defined in Chapter 2 and Appendix A.

Table A3-1. Three-digit Recodes--Diseases of the Respiratory System

---

144.	Common Cold *
145.	Other Acute Upper Respiratory Infections *
146.	Acute Bronchitis *
147.	Influenza with Digestive Manifestations *
148.	Other Forms of Influenza *
149.	Viral Pneumonia *
150.	Other Forms of Pneumonia *
151.	Chronic Bronchitis (nonallergic)
152.	Emphysema
153.	Asthma (with or without hay fever) (allergic)
154.	Hypertrophy of Tonsils and Adenoids, Chronic
155.	Chronic Pharyngitis, Nasopharyngitis, and Laryngitis, NEC
156.	Chronic Sinusitis, NEC
157.	Hay Fever, and Upper Respiratory Allergy, without Asthma
158.	Other Diseases of Upper Respiratory Tract (nonallergic)
159.	Other Diseases of Respiratory System, NEC
228.	Certain Respiratory Symptoms *

---

* Can be coded as acute if onset is less than three months prior.

The respiratory condition measures are:

7) TRADRSP - This variable is defined as the total number of two-week restricted-activity days due to respiratory conditions. Specifically, for i indexing over {01, 02, ... , 14},

$$\text{TRADRSP} = \text{Max}_i\{\ \text{RAD2W}_i \mid \text{DGRCD1}_i \text{ in Table A3-1 above } \}$$

That is, TRADRSP is the maximum of the set of two-week restricted-activity days given that such days were due to a respiratory condition. If no respiratory-related two-week restricted-activity days were reported, then TRADRSP = 0.

8) TWSLRSP - This variable is defined as the total number of two-week work- or school-loss days due to respiratory conditions. Specifically, again for i indexing over {01, 02, ... , 14 },

For Adults: $\text{WSLRESP} = \text{Max}_i\{\ \text{WSL2W}_i \mid \text{DGRCD1}_i \text{ in Table A3-1 above } \}$

if CURACT2W = 1 or CURACT2W = 2, and n/a otherwise.

For Children:

$$\text{WSLRESP} = \text{Max}_i\{\ \text{WSL2W}_i \mid \text{DGRCD1}_i \text{ in Table A3-1 above } \}$$

if WKSLLDID = 2, and n/a otherwise.

The maximum here, of course, can be zero for those individuals reporting no respiratory-related work- or school-loss days.

9) TBEDRSP - This variable is defined as the total number of two-week bed-disability days due to respiratory conditions. Specifically, with i again taking on values 01, ... , 14,

$$\text{BEDRESP} = \text{Max}_i\{\ \text{BED2W}_i \mid \text{DGRCD1}_i \text{ in Table A3-1 above } \}$$

Again, no respiratory-related bed-disability days implies BEDRESP = 0.

10) RADRSPPS - This is a binary variable, constructed from TRADRSP as follows:

RADRSPPS = 1 iff TRADRSP > 0
RADRSPPS = 0 iff TRADRSP = 0.

11) WSLRSPPS - This is a binary variable, derived from WSLRESP as follows:

WSLRSPPS = 1 iff WSLRESP > 0
WSLRSPPS = 0 iff WSLRESP = 0.

12) BEDRSPPS - This is a binary variable, constructed from BEDRESP as follows:

BEDRSPPS = 1 iff BEDRESP > 0
BEDRSPPS = 0 iff BEDRESP = 0.

13) PRSCHRSP - This is a measure of chronic respiratory illness. Because of the nature of such chronic conditions, a discrete binary measure was deemed appropriate. PRSCHRSP is defined as

PRSCHRSP = 1 iff, for any i in { 01, ... , 14 },
DGRCD1_i is in Table A3-1 above and CRNACT_i = 1
PRSCHRSP = 0 otherwise.

Thus PRSCHRSP assumes a value of one when the individual reports any respiratory condition listed in Table A3-1 defined by the standards of the HIS to be chronic, and zero otherwise.

Preliminary explorations into the possible effects of air pollution on cardiovascular illness have been undertaken in this study. For present purposes, cardiovascular illnesses were classified as those values of DGRCD101 through DGRCD114 listed in Table A3-2 below.

Table A3-2. Three-digit Recodes--Selected Diseases of the Circulatory System

---

128.	Chronic Rheumatic Heart Diseases
129.	Ischemic Heart Diseases (with Hypertension, Any Type) (with Arteriosclerosis)
130.	Heart Trouble, NOS, or Ill-defined
131.	Other Forms of Heart Disease, NEC
132.	Hypertensive Heart Disease, NEC (Non-malignant)
133.	Hypertensive Disease, NEC
134.	Cerebrovascular Disease (with Non-malignant Hypertension) (with Arteriosclerosis)
135.	Arteriosclerosis, NEC
136.	Other Diseases of the Arteries
141.	Poor Circulation, NOS
143.	Other Diseases of Circulatory System, NEC

The single measure of cardiovascular conditions used was PRSCHCRD, defined as:

$$\text{PRSCHCRD} = 1 \text{ iff, for any } i \text{ in } \{01, \ldots, 14\},$$
$$\text{DGRCD1}_i \text{ is in Table A3-2 above } \underline{\text{and}} \text{ CRNACT}_i = 1$$
$$\text{PRSCHCRD} = 0 \text{ otherwise.}$$

Thus, PRSCHCRD assumes a value of one when the individual reports any cardiovascular condition listed in Table A3-2 and defined by the standards of the HIS to be chronic, and zero otherwise.

Chapter 4

RESULTS AND DISCUSSION

In this chapter we report our findings about acute and chronic morbidity and their relationship to ozone and other air pollutants. First, some introductory remarks are in order.

In Chapter 1, several problems inherent in epidemiological studies were discussed. Other problems arise in attempts to estimate dose-response relationships via statistical regression. First, this approach requires one to specify a functional form for the dose-response relationship, yet there is very little theoretical guidance as to its likely nature. Some studies use a simple linear form, others rely on a logarithmic specification, while many recent studies have employed an S-shaped probit or logistic functional form. There are other possible choices, as well, with examples of each in the literature.

Nor is there consensus about the "target populations" to be studied or about the sorts of adverse health effects on which to concentrate attention. One can concentrate on acute or chronic morbidity; examine respiratory as opposed to other effects; focus on the very young or the very old, or the general population; concentrate on the healthy as opposed to the infirm; and so on.

In addition, there is disagreement about the explanatory variables that should be included, or how they should be expressed, as for example whether air pollutants should be measured by peak or average concentrations. Some studies use a single air pollutant to characterize exposure to all pollutants, some include few socioeconomic or meteorological measures, some fail or are unable to control for the smoking habits of individuals. In most studies, as in the present one, the number of control variables is determined as much by the availability of information as by the conscious choice of the researcher to include or exclude certain measures.

Together, these factors--the range of possible and plausible functional forms, types of health effects to be studied, choice of target populations, and variety of explanatory variables to consider for inclusion--have important implications first for the design or conduct of research and second for the presentation and interpretation of results. Concerning the former, they suggest that, where possible, one should experiment with a variety of functional forms, types of health effects, target populations and control measures.

That is the approach taken in this report. Where a choice had to be made, we elected to explore a broad range of possible associations between air pollution and acute and chronic disease using fairly standard ordinary least squares and logit techniques. Because of the several unique features of our data, it seemed advisable to take this route rather than confine our attention to a few relationships which would then be subjected to a battery of sophisticated tests. Although the approach in this report is designed to help identify "targets of opportunity" for further research, the relationships we estimate do not go unexamined. The results to follow include a great many sensitivity analyses, and sophisticated regression diagnostic procedures have been applied to several of the most interesting or important findings.

With respect to the presentation of results, such an approach means it may be virtually impossible to present in a report every model that was estimated. In a study like the present one the number of separate models explored often extends into the hundreds, thus forcing a decision about which results to present. In making this decision, it is important to convey the full range of work done and results obtained. This is to be contrasted with an approach in which only the "best" results are displayed, that is, those which enable one to either reject or accept the null hypothesis of no relationship between air pollution and morbidity. Selective reporting can give a misleading picture of this relationship.

The presentation of results in this report is designed to minimize this problem. The results presented here represent fairly well the range of functional forms, categories of health effects, target populations and combinations of explanatory variables explored in the study. Generally, within a particular area of inquiry--acute respiratory disease in adults, for instance--several versions of a "basic" model are presented. These versions may differ with respect to the number of air pollutants besides ozone which are analyzed, the way these pollutants are measured, and other dimensions.

Our approach could be termed conservative in several respects. First, time and financial constraints made it impossible to analyze the sensitivity of every finding. We elected to do much more sensitivity analysis for those models in which the null hypothesis of no effect of ozone on acute or chronic morbidity was rejected at standard confidence levels. The goal was to explore the observed effect for statistical robustness. In some models where no statistically significant effect was observed, little or no additional analysis was done. In certain cases, however, sensitivity tests were performed even if no significant effect was observed.

Another respect in which our approach might be termed conservative concerns hypothesis testing. Throughout the results, two-tailed tests are used to determine the significance of the coefficients on ozone and other explanatory variables. Using two-tailed tests, t-ratios on estimated coefficients must be $\geq$ 1.96 for the coefficient to be significant at the 5 percent level and $\geq$ 1.65 for significance at the 10 percent level.

This approach might be termed conservative because some epidemiological studies have used one-tailed tests on the grounds that if air pollution has any effect on health, the effect will only be adverse.[1] When using a one-tailed test, the t-ratio on an estimated coefficient need only be 1.65 for the coefficient to be significant at the 5 percent level and 1.28 for significance at the 10 percent level. The more rigorous requirement associated with the two-tailed test is used here.

The parameter estimates and t-statistics presented below are computed using the usual formulae:

$$\hat{\beta} = (X^TX)^{-1}X^TY \quad \text{and} \quad t_i = \hat{\beta}_i \text{SQRT}\ [\text{DIAG}(S^2(X^TX)^{-1})_i].$$

For each OLS regression, we present at the bottom of the column of results the uncorrected R^2, the F-statistic testing the hypothesis H_0: $\beta_{-1} = 0$ (where β_{-1} is the parameter vector except for the estimated intercept), and the number of observations in the estimated equation.

For logit models, we present the maximum likelihood estimate of each parameter ($\hat{\beta}_i$) and the estimated t-statistic. The latter is $\hat{\beta}_i$ divided by the square root of the corresponding diagonal element of I^{-1} evaluated at the maximum likelihood parameter estimates (I is the information matrix). At the bottom of the columns presenting logit results are found the χ^2 statistics (comparable to F-statistics). χ^2 is calculated as $-2(\ln L_R - \ln L_F)$ where L_R and L_F are, respectively, the likelihood ratios for the intercept only, and for the model with all parameter estimates. We also present there the number of observations used to estimate the equation.

4.1 Acute Morbidity Due to All Causes

4.1.1 Adults

The first results concern bed disability days, work loss days, and minor restricted activity days (MRADs) during the respondent's two-week reference period--the latter being days on which individuals are not ill enough to stay in bed or miss work, but must restrict their usual activity. Thus, a MRAD may be considered as the slightest manifestation of acute morbidity used in our study. Unfortunately, the HIS makes it difficult to identify these minor restrictions because any work loss or bed disability day is also coded as a restricted activity day. To be sure we were working with truly minor restrictions, we confined this analysis to only those individuals reporting RADs who did not also have work loss or bed disability days. The sample thus excludes those who reported more serious acute illnesses during the reference period. The acute morbidity analyzed here was that due to all causes, not just respiratory.

The first regressions, presented in Table 4-1, used ordinary least squares to relate MRADs to ozone (as measured by the average daily maximum 1-hour ozone concentration over the two-week period), other air pollutants, and additional explanatory variables. In these and all other results presented below air pollution was characterized by the readings at the monitor nearest each individual's house for which data were available. No individuals were included in any of the statistical analyses in this report if the nearest monitor with valid data was greater than twenty miles from their home.

In equation (1), the only air pollutant included besides ozone was a measure of sulfate concentrations during the two-week period. We concentrated most heavily on these two pollutants throughout this report because ozone and sulfates (particularly highly acidic sulfate particles) appear to be the two pollutants most likely to be harmful at levels frequently observed in the United States today. However, considerable attention was also devoted to total suspended particulate matter, carbon monoxide, nitrogen dioxide and sulfur dioxide in the analysis. Ozone was

Table 4-1. Restricted Activity Days Among Adults: All Causes

Model #	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dep. Var.										
Indep. Var.	RADMINOR	RADMINOR	RADMINOR	RADMINOR	RADMINOR	RADMINOR	RADMINOR	RADMINOR	RADPRESM (LOGIT)	RADRESSM (LOGIT)
INTERCEPT	.772 (1.21)	1.129 (2.14)	.540 (0.77)	1.352 (1.75)	.312 (1.27)	1.158 (2.19)	.689 (1.07)	1.412 (1.82)	-1.595 (1.10)	-1.543 (1.07)
D3NR01	2.428 (1.73)	1.489 (1.49)	2.093 (1.54)	1.642 (1.11)	--	--	--	--	6.119 (2.17)	--
D3NR01SQ	--	--	--	--	14.489 (2.48)	9.910 (1.81)	14.878 (2.18)	12.554 (1.67)	--	31.052 (2.43)
S4NR01	-.0133 (2.44)	--	-.0150 (2.73)	-.0230 (3.56)	-.0131 (2.51)	--	-.0148 (2.99)	-.0229 (3.56)	-.0175 (0.20)	-.0162 (0.10)
SPNR01	--	.00134 (1.95)	.00150 (1.53)	.00196 (1.69)	--	.00129 (1.85)	.00139 (1.51)	.00174 (1.57)	.000434 (1.30)	.000230 (1.20)
N2NR01	--	--	--	.0000172 (0.02)	--	--	--	-.000454 (0.06)	--	--
CONR01	--	--	--	.00146 (0.15)	--	--	--	.00213 (0.17)	--	--
RACEW130	-.00136 (0.02)	-.019 (0.20)	-.000341 (0.01)	-.106 (1.04)	-.000135 (0.00)	-.0187 (0.27)	.000263 (0.00)	-.105 (1.05)	-.101 (0.52)	-.0944 (0.49)
SEXM1F0	-.0265 (0.44)	-.0354 (0.74)	-.0270 (0.45)	-.000134 (0.00)	-.0272 (0.45)	-.0364 (0.75	-.0275 (0.46)	-.000624 (0.01)	-.300 (2.07)	-.300 (2.07)
MARYIN0	.0228 (0.36)	.0659 (1.29)	.0225 (0.36)	.0434 (0.65)	.0241 (0.39)	.0655 (1.30)	.0237 (0.38)	.0499 (0.67)	.0557 (0.36)	.0594 (0.39)
INCOMCON	-.00000469 (2.64)	-.0000112 (4.21)	-.00000334 (2.53)	-.00000936 (2.55)	-.00000372 (2.58)	-.0000112 (4.25)	-.00000840 (2.55)	-.0000101 (2.59)	-.0000237 (2.80)	-.0000243 (2.84)
CITYI100	-.0141 (0.25)	-.0265 (0.51)	-.0205 (0.36)	-.110 (1.92)	-.0124 (0.22)	-.0263 (0.53)	-.0184 (0.33)	-.127 (1.89)	.0439 (0.30)	.0430 (0.30)
FAT	-.303 (0.56)	-.716 (1.93)	-.282 (0.63)	-.630 (1.17)	-.304 (0.60)	-.716 (1.92)	-.294 (0.63)	-.629 (1.17)	-.861 (0.94)	-.871 (0.94)
FATSQ	.0574 (0.56)	.137 (1.83)	.0537 (0.52)	.119 (1.14)	.0575 (0.56)	.137 (1.84)	.0340 (0.52)	.119 (1.12)	.185 (1.06)	.187 (1.07)

Reproduced from best available copy.

Table 4-1 (Continued). Restricted Activity Days Among Adults: All Causes

Model #	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dep. Var. / Indep. Var.	RADMINOR	RADMINOR	RADMINOR	RADMINOR	RADMINOR	RADMINOR	RADMINOR	RADMINOR	RADPRESM (LOGIT)	RADRESSM (LOGIT)
AGE	.00313	-.00545	.00370	.0174	.00393	-.00550	.00354	.0174	.0271	.0272
	(1.00)	(.073)	(0.35)	(1.61)	(0.38)	(0.74)	(0.93)	(1.60)	(1.24)	(1.24)
AGESQ	-.0000691	.000113	-.0000549	-.000160	-.0000567	.000113	-.0000630	-.000159	-.000253	-.000260
	(0.72)	(1.43)	(0.57)	(1.40)	(0.69)	(1.44)	(0.65)	(1.40)	(1.20)	(1.20)
SMOKYIND	.149	.0720	.149	.143	.148	.0719	.147	.142	.408	.405
	(2.63)	(1.56)	(2.50)	(2.11)	(2.52)	(1.55)	(2.59)	(2.10)	(2.91)	(2.89)
EDCOMCON	-.00734	-.000316	-.00671	-.0114	-.00921	-.000354	-.00664	-.0112	-.0225	-.0223
	(0.76)	(0.04)	(0.70)	(1.01)	(0.75)	(0.05)	(0.69)	(0.99)	(0.99)	(0.99)
CHRLMDUM	1.297	1.451	1.293	1.341	1.295	1.460	1.297	1.339	1.305	1.899
	(16.37)	(23.10)	(15.36)	(14.23)	(16.35)	(23.09)	(16.34)	(14.27)	(12.92)	(12.86)
CROWDING	-.112	-.139	-.119	-.104	-.109	-.138	-.115	-.101	-.623	-.612
	(1.21)	(1.82)	(1.27)	(0.99)	(1.18)	(1.81)	(1.24)	(0.96)	(2.33)	(2.29)
AVMAXTMP	.000235	-.000528	.000170	-.000199	.000957	-.000141	.000685	.000192	-.00495	-.00234
	(0.12)	(0.33)	(0.09)	(0.03)	(0.49)	(0.10)	(0.39)	(0.03)	(1.03)	(0.56)
AVPRECIP	.414	-.289	.501	.166	.441	-.284	.521	.194	1.359	1.341
	(1.49)	(1.31)	(1.77)	(0.43)	(1.59)	(1.30)	(1.84)	(0.56)	(1.39)	(1.98)
HUMIDRF	-.00413	.00169	-.00366	-.00550	-.00452	.00149	-.00407	-.00601	-.00749	-.00805
	(1.21)	(0.64)	(1.04)	(1.26)	(1.32)	(0.56)	(1.15)	(1.38)	(0.97)	(0.93)
R2	.0783	.0938	.0797	.0827	.0793	.0939	.0901	.0831		
F or X2	22.53	45.32	21.55	16.30	22.69	46.39	21.67	15.38	292.52	293.34
N	4,759	8,176	4,751	3,534	4,759	8,175	4,751	3,594	4,751	4,751

positively related to the number of MRADs during the two-week period, a relationship significant at the 10 percent level. Surprisingly, the sulfate variable was negatively and significantly related to MRADs. We have no explanation for this finding.

Of the other variables in equation (1), only three were significant. They were income (INCOMCON), the smoking dummy, and the dummy variable indicating the presence of a chronic illness or impairment. All three of these variables had the expected signs, suggesting that higher income means fewer days of acute illness, while smoking and the presence of a chronic illness imply more acute illness. The t-ratio on the chronic illness dummy demonstrates the very strong effect it exerts on the number of MRADs during a two-week period. The R^2 for the regression is 0.08 and the F-statistic is 22.5.

Equations (2) - (4) replicate this regression but with different combinations of air pollution variables. In equation (2) sulfates were replaced by total suspended particulates. The coefficient on ozone remained positive (as it did throughout virtually all of the analysis) although it was no longer significant at the 10 percent level. Particulates, on the other hand, were positively related to MRADs, at the 10 percent level. The coefficients on the chronic limitation and income variables increased in significance, while that on the smoking dummy fell. Additionally, the coefficients on both FAT and FATSQ (the proxy for general physical condition) increased in significance. Their signs suggest a U-shaped relationship between weight-to-height ratio and acute morbidity.

In equation (3), ozone was entered with both sulfate concentrations as well as total suspended particulates. Ozone again was positive although not significant at the 10 percent level. Particulates were positively associated with acute illness while sulfates had a negative association significant at the 5 percent level. The coefficient on the smoking variable was significant at the 5 percent level. Equation (4) included five different pollutants. Of the five, only sulfates were significant at the 5 percent level, again, however, with an unexpected sign; particulate matter was positively associated with MRADs at the 10 percent level. The coefficient estimates on the other independent variables in (3) and (4) were similar to those in equation (1).

In equations (1) - (4) the number of individuals included in a particular regression varied from 3,583 to 8,175. This is because when more air pollution data were required for a particular regression, fewer individuals were likely to have them. The largest sample size occurred in regressions with ozone and particulate matter because of the large number of particulate monitors around the country compared to the other pollutants.

Equations (5) - (8) replicate (1)- (4) with one important difference. Since the relationship between ozone and acute morbidity may be non-linear (increments to the average daily maximum 1-hour ozone concentration may do more harm at higher than at lower levels), we reexamined the determinants of MRADs using the square of the average daily maximum 1-hour concentration to measure ozone during the two-week reference period.[2] As (5) illustrates, the coefficient on the square of ozone was of the expected sign and was significant at the 1 percent level. Once again, income, smoking and chronic illness were the other variables significantly associated with acute morbidity as measured by MRADs. In equations (6) - (8) ozone was always significant at the 5 or 10 percent level.

In equations (9) and (10), this basic model was reexamined again, this time using a logit specification rather than ordinary least squares. [3] Any individual having at least one MRAD due to any cause during the two-week period was given the value unity, while those with no MRADs were assigned the value zero. (RADPRESM connotes the presence of at least one MRAD.) Equation (9) relates in a linear fashion the log of the odds that an individual had at least one day of acute illness to the set of explanatory variables. As the t-ratio indicates, ozone was positively and significantly (at the 5 percent level) related to the likelihood of experiencing a MRAD. The regression suggests that women are more likely than men to experience at least one MRAD, as are smokers, those with lower incomes, individuals with a chronic illness, and those living in areas where precipitation is relatively high. The measure of persons per room in the household was also significant at the 5 percent level although the sign is not as expected. Equation (10) replicates (9) using the square of the average daily maximum 1-hour ozone concentration. As in regressions (5) - (8), this change increased the significance of the coefficient on ozone, in this case nearly to the 1 percent level. In both (9) and (10) the coefficient on sulfates, while still negative, was no longer significant.

Table 4-2 presents the analysis of work loss and bed disability in adults due to all causes. Equations (11) - (13) experiment with different

Table 4-2. Work Loss and Bed Disability Among Adults: All Causes

Model #	(11)	(12)	(13)	(14)	(15)	(16)
Dep. Var. Indep. Var.	WKSLLD2W	WKSLLD2W	WKSLLD2W	BEDDIS2W	BEDDIS2W	BEDDIS2W
INTERCEPT	.403 (0.73)	.521 (0.94)	.562 (0.86)	1.065 (2.13)	1.178 (2.33)	1.844 (3.21)
O3NR01	-.691 (0.65)	-.448 (0.42)	-.671 (0.55)	-.498 (0.49)	-.248 (0.24)	-.419 (0.37)
S4NR01	-.0014 (0.33)	-.000245 (0.06)	-.00222 (0.40)	.00238 (0.60)	.00361 (0.88)	-.000301 (0.06)
SPNR01	-	-.00114 (1.41)	-.000706 (0.74)	--	-.00118 (1.59)	-.00138 (1.62)
N2NR01	--	--	-.000332 (0.56)	--	--	.000652 (1.19)
CONR01	--	--	.00760 (0.78)	--	--	.00154 (1.74)
RACEW1B0	.0462 (0.65)	.0457 (0.64)	.0758 (0.92)	.115 (1.68)	.115 (1.68)	.129 (1.68)
SEXM1F0	-.0988 (1.88)	-.0976 (1.85)	-.158 (2.55)	-.115 (2.40)	-.114 (2.37)	-.0852 (1.56)
MARY1N0	.0411 (0.77)	.0418 (0.78)	.0501 (0.80)	-.0169 (0.33)	-.0162 (0.32)	-.0338 (0.59)
INCOMCON	-7.336E-7 (0.36)	-9.952E-7 (0.35)	-4.323E-7 (0.13)	-.00000612 (2.33)	-.00000637 (2.42)	-.00000357 (1.20)
CITYI100	.0172 (0.37)	.0197 (0.42)	.000277 (0.01)	.0563 (1.25)	.0586 (1.30)	.106 (2.04)
FAT	-.273 (0.72)	-.287 (0.76)	-.251 (.057)	-.813 (2.36)	-.823 (2.38)	-1.510 (3.87)
FATSQ	.0632 (0.87)	.0659 (0.91)	.0731 (0.86)	.203 (3.06)	.205 (3.08)	.345 (4.57)
AGE	.0144 (1.39)	.0146 (1.41)	.00614 (0.51)	-.00782 (1.07)	-.00759 (1.04)	-.00797 (0.97)

Table 4-2 (Continued). Work Loss and Bed Disability Among Adults: All Causes

Model #	(11)	(12)	(13)	(14)	(15)	(16)
Dep. Var. Indep. Var.	WKSLLD2W	WKSLLD2W	WKSLLD2W	BEDDIS2W	BEDDIS2W	BEDDIS2W
AGESQ	-.000221 (1.82)	-.000223 (1.82)	-.000127 (0.89)	.0000758 (0.99)	.0000737 (0.96)	.0000934 (1.08)
SMOKY1NO	-.0164 (0.35)	-.0149 (0.32)	-.0141 (0.28)	.0614 (1.36)	.0625 (1.37)	.0272 (0.52)
EDCOMCON	-.0204 (2.24)	-.0205 (2.25)	-.0250 (2.35)	-.00985 (1.28)	-.0102 (1.32)	-.0150 (1.74)
CHRLMDUM	.396 (4.78)	.398 (4.80)	.335 (3.47)	.716 (11.71)	.719 (11.74)	.733 (10.59)
CROWDING	-.0738 (0.94)	-.0658 (0.84)	-.133 (1.51)	.00219 (0.03)	.00693 (0.09)	-.00292 (0.04)
AVMAXTMP	.00104 (0.63)	.00113 (0.68)	.00272 (1.38)	-.00199 (1.25)	-.00191 (1.20)	-.000930 (0.51)
AVPRECIP	-.214 (0.93)	-.274 (1.17)	-.290 (1.04)	-.304 (1.36)	-.373 (1.64)	-.0505 (0.19)
HUMIDRF	.00461 (1.63)	.00381 (1.31)	.00397 (1.12)	.00561 (2.05)	.00488 (1.75)	.00381 (1.15)
INDOCCAV	-.0202 (2.38)	-.0207 (2.41)	-.0185 (1.83)	--	--	--
R^2	.0186	.0193	.0210	.0483	.0488	.0589
F or X^2	2.83	2.78	2.04	14.50	13.96	11.61
N	2,854	2,851	2,107	5,200	5,193	3,918

combinations of air pollutants for adult work loss, while similar combinations are used in (14) - (16) to explain adult bed disability during the two-week reference period. As indicated, none of the air pollution variables was significant at the 5 percent level in any of the regressions, and only carbon monoxide (CONR01) was significant at the 10 percent level--in the regression explaining adult bed disability. The R^2's were much lower in these equations than those explaining MRADs. Substituting the square of ozone made no difference in any of these results. Of the other explanatory variables, the chronic illness dummy was still the most important. It was highly significant in all the equations. Equations (11) - (13) suggest that women experience more work loss and bed disability days than men, and that the amount of work loss is negatively and significantly related to education.

Of particular interest in the work loss regressions is the measure of paid sick leave, INDOCCAV. It was significant at the 5 percent level in two of the equations and at the 10 percent level in the third. Its sign suggests that work loss varies inversely with paid sick leave. This may reflect the willingness of workers with ample paid sick leave to miss an occasional day of work in the early stages of illness, thus preventing prolonged absences due to illnesses that are allowed to grow more serious. It may also reflect the fact that the least physically demanding jobs (in offices, for example) often carry with them generous paid sick leave. Thus, the measure of paid sick leave may actually be distinguishing between relatively safe and "clean" jobs and those where hazards may be present. To explore this possibility, an additional work loss model was estimated (not presented here) in which a blue collar dummy variable was substituted for INDOCCAV. The dummy variable was positively related to work loss and was significant at the 1 percent level, thus giving support to the latter suspicion.

Because of the positive and often statistically significant association between ozone and MRADs due to all causes in adults, a variety of sensitivity analyses were performed to determine the robustness of the relationship. These are presented in Table 4-3. In equation (17), attention is concentrated on a subset of individuals for whom data are available on the usual set of regressors, but also for ragweed (pollen) concentrations.[4] Equation (17) is identical to equation (1) save for the pollen variable. The introduction of this variable (i) reduced the magnitude of the coefficient of the ozone variable; (ii) reduced the t-ratio on the ozone coefficient from 1.93 to 1.12; and (iii) reduced the sample size by more than half because pollen data were only available for a subset of the full sample.

In (18) - (20) ozone was measured not by the average daily maximum 1-hour concentration over the two-week period but rather by the average hourly concentration over that same period. (Rather than average the 14 daily highest one-hour readings, all hourly readings up to a maximum of 336 were averaged.) This change reduced the significance of the ozone variable when compared to equations (1) - (4). This is not surprising, however, since it is commonly thought that peak rather than average ozone concentrations are most likely to induce or exacerbate acute morbidity.

Table 4-3. Sensitivity Analyses of Restricted Activity in Adults: All Causes

Model #	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)	(27)
Dep. Var. / Indep. Var.	RADMINOR	RADMINOR	RADMINOR	RADMINOR	RADMINOR	RADMINOR	RADMINOR	RADMINOR	RADMINOR	RADMINOR	RADMINOR
INTERCEPT	.803 (0.81)	.718 (1.13)	.579 (0.90)	1.217 (1.64)	.435 (0.53)	1.417 (2.02)	.245 (0.36)	1.177 (1.21)	.480 (0.74)	1.112 (1.69)	.654 (1.91)
O3NR01	1.459 (1.12)	--	--	--	1.595 (0.37)	1.524 (1.15)	1.252 (0.75)	.832 (0.46)	5.211 (2.97)	-14.390 (2.19)	--
O3NR04	--	2.729 (1.03)	2.334 (0.38)	.814 (0.21)	--	--	--	--	--	--	--
O3S4	--	--	--	--	--	--	--	--	-.279 (2.48)	--	--
O3AVMAXT	--	--	--	--	--	--	--	--	--	.227 (2.61)	--
O3H15	--	--	--	--	--	--	--	--	--	--	3.620 (2.37)
O3L15	--	--	--	--	--	--	--	--	--	--	.305 (0.19)
S4NR01	-.0167 (2.21)	-.0126 (2.51)	-0145 (2.35)	-.0224 (3.47)	-.0141 (2.13)	-- (2.37)	-.0164 (2.37)	-.0245 (2.34)	--	-.0131 (2.49)	-.0142 (2.75)
SPNR01	--	--	.00167 (1.43)	.00202 (1.93)	--	.00190 (1.91)	.00161 (1.32)	.00150 (1.05)	.00139 (1.51)	.00206 (2.21)	.00136 (1.47)
N2NR01	--	--	--	.000200 (0.29)	--	--	--	.000171 (0.11)	--	--	--
CONR01	--	--	--	.000810 (0.07)	--	--	--	.00741 (0.43)	--	--	--
RACEWH10	-.139 (1.17)	-.00136 (0.02)	-.000501 ([illegible])	-.108 (1.60)	[illegible]	.00151 (0.02)	.0311 (0.33)	.0337 (0.31)	.00322 (0.11)	-.0218 (0.25)	.00181 (0.02)
SEXM2F0	.0227 (0.25)	-.0247 (0.41)	-.0263 (0.63)	.0114 (0.07)	[illegible]	.0220 (0.33)	.0210 (0.13)	.0690 (0.44)	-.0266 (0.44)	-.00720 (0.12)	-.0272 (0.45)

Table 4-3 (Continued). Sensitivity Analyses of Restricted Activity in Adults: All Causes

Model #	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)	(27)
Dep. Var. / Indep. Var.	RADMINOR	RADMINOR	RADMINOR	RADMINOR	RADMINOR	RADMINOR	RADMINOR	RADMINOR	RADMINOR	RADMINOR	RADMINOR
MARYIND	.116 (1.21)	.0239 (.033)	.0231 (0.37)	.0514 (0.63)	.0752 (0.91)	.0854 (1.24)	.0762 (0.92)	.0997 (1.05)	.0250 (0.40)	.0230 (0.45)	.0250 (0.40)
INCOMCON	-.0000144 (2.93)	-.00000858 (2.64)	-.00000930 (2.52)	-.00000991 (2.54)	-.0000121 (2.61)	-.0000127 (3.35)	-.0000118 (2.54)	-.0000144 (2.69)	-.00000950 (2.58)	-.00000772 (2.34)	-.00000840 (2.55)
CITYIIUO	-.165 (1.72)	-.0190 (0.32)	-.0242 (0.43)	-.136 (2.01)	-.0243 (0.33)	-.0524 (0.85)	-.0304 (0.40)	-.173 (1.96)	-.0245 (0.44)	-.0194 (0.33)	-.0180 (0.32)
FAT	-.243 (0.34)	-.303 (0.69)	-.280 (0.52)	-.631 (1.17)	-.356 (0.64)	-.946 (1.95)	-.334 (0.50)	-.752 (1.15)	-.289 (0.65)	-.242 (0.59)	-.279 (0.62)
FATSQ	.0343 (0.24)	0.571 (0.65)	.0331 (0.51)	.119 (1.14)	.0714 (0.56)	.183 (1.94)	.0474 (0.42)	.142 (1.12)	.0552 (0.63)	.0436 (0.57)	.0531 (0.61)
AGE	.0114 (0.84	.00927 (1.01)	.00373 (0.96)	.0115 (1.61)	.0101 (0.81)	-.00706 (0.63)	.00360 (0.77)	.0137 (1.23)	.00875 (0.96)	.00998 (1.07)	.00950 (0.93)
AGESQ	-.0000535 (0.38)	-.0000706 (0.73)	-.0000657 (0.48)	-.000151 (1.41)	-.0000706 (0.55)	.000134 (1.25)	-.0000553 (0.50)	-.000157 (1.04)	-.0000653 (0.68)	-.000859 (0.88)	-.0000626 (0.45)
SMOKYINO	.173 (2.07)	.148 (2.62)	.147 (2.59)	.142 (2.11)	.193 (2.42)	.0836 (1.33)	.196 (2.58)	.190 (2.05)	.149 (2.62)	--	.146 (2.58)
NCIGSDYN	--	--	--	--	--	--	--	--	--	.00155 (0.69)	--
EDCOMCON	-.0114 (0.41)	-.00732 (0.75)	-.00566 (0.45)	-.0114 (1.01)	-.00201 (0.16)	.00230 (0.72)	-.000795 (0.06)	-.00323 (0.22)	-.00656 (0.68)	-.0123 (1.27)	-.00660 (0.58)
CHRIMDUM	1.389 (11.42)	1.298 (15.37)	1.293 (14.36)	1.251 (14.27)	1.453 (14.72)	1.654 (20.14)	1.489 (14.69)	1.494 (12.73)	1.297 (16.33)	1.261 (15.59)	1.295 (16.32)
POLLENRF	.0121 (0.37)	--	--	--	--	--	--	--	--	--	--
CROWDING	-.0212 (0.15)	-.105 (1.17)	-.117 (1.26)	-.192 (0.91)	-.109 (0.97)	-.133 (1.34)	-.112 (0.30)	-.0638 (0.510	-.117 (1.26)	-.132 (1.410	-.114 (1.23)

Reproduced from best available copy.

Table 4-3 (Continued). Sensitivity Analyses of Restricted Activity in Adults: All Causes

Model #	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)	(27)
Dep. Var.	RADMINOR	RADMINOR	RADMINOR	RADMINOR	RADMINOR	RADMINOR	RADMINOR	RADMINOR	RADMINOR	RADMINOR	RADMINOR
Indep. Var.											
AVMAXTMP	.0000374 (0.03)	.00142 (0.74)	.00115 (0.50)	.000971 (0.42)	.00145 (0.55)	-.00198 (0.50)	.00140 (0.53)	.000132 (0.04)	.000366 (0.18)	-.00638 (2.13)	.00150 (0.70)
AVPRECIP	.226 (0.48)	.347 (1.26)	.454 (1.61)	.117 (0.34)	.490 (1.32)	-.138 (0.45)	.795 (2.06)	.167 (0.36)	.486 (1.72)	.551 (1.93)	.494 (1.74)
HUMIDRF	-.00127 (0.25)	-.00359 (1.09)	-.00317 (0.91)	-.00516 (1.11)	-.00212 (0.47)	.000779 (0.22)	-.00171 (0.37)	-.00322 (0.57)	-.00374 (1.07)	-.00250 (0.71)	-.00418 (1.19)
R2	.0343	.0733	.0793	.0825	.0913	.1018	.0921	.0946	.0795	.0765	.0803
F or X2	12.35	22.38	21.45	15.24	17.85	34.59	17.05	12.04	21.54	19.10	20.65
N	2,274	4,758	4,251	3,534	3,217	5,512	3,213	2,441	4,751	4,632	4,751

Equations (21) - (24) replicate equations (1) - (4) with one difference. The sample of individuals in (21) - (24) were what we term "self-respondents." That is, during the HIS interview these individuals were present and answered the interviewer's questions for themselves (if an adult was not present during the HIS, his or her spouse--or other available adult--provided the relevant information for the absentee). Self-respondents were singled out on the presumption that information about their acute morbidity might be more accurate than if another had provided information for or about them. Across the four regressions, the coefficients on the ozone variable were generally smaller and less significant than in the larger sample that included both self-respondents and those for whom information was provided by others. This casts some doubt on the earlier positive and significant associations between ozone and MRADs. We cannot rule out the possibility that the smaller and less significant coefficients are due to the smaller sample size, however. Other sensitivity analyses presented later in this chapter make use of the distinction between self-respondents and others.

Possible interactive or synergistic effects of ozone with other pollutants or meteorological variables were also explored. In equation (25), both the average daily maximum 1-hour ozone concentration and a multiplicatively interactive term involving ozone and sulfates were used as explanatory variables along with particulates and the other variables in the basic model.[5] By itself, ozone was positively and significantly related to MRADs. However, the sign of the interactive term was negative and significant. In equation (26) ozone was allowed to interact with the average daily maximum temperature. Although the coefficient on ozone in this model was negative and significant, the interactive term entered positively and significantly.

With the exception of the logit regressions and those using the square of the ozone concentration, the analysis to this point has been based on linear regression over the whole range of ozone concentrations. Yet many clinical studies are designed to determine whether some "threshold" or lowest observed effect ozone concentration exists, below which health effects are not detected.[6] Accordingly, several regressions were run using "switching" techniques which permit different relationships to be identified for different ranges of ozone.[7] Equation (27), which had the highest F-statistic of the switching regressions we tried, presents results from a regression where different relationships between ozone and acute morbidity could be identified for average daily maximum 1-hour concentrations above (03H15) and below (03L15) 0.15 parts-per-million (ppm).

In (27), for average daily maximum 1-hour ozone concentrations greater than 0.15 ppm, the coefficient on the ozone variable was 3.62 (the slope measure) and it was significant at the 5 percent level. Below 0.15 ppm, however, the slope was much flatter (0.3) and the coefficient was not significantly different from zero. This suggests that a different relationship may exist between ozone and MRADs from all causes above and below that level, a finding consistent with clinical research relating ozone exposure to lung function.[7a] Regressions were run using different switch points both greater and less than 0.15 ppm but in no case was there such a clear difference in slope and significance between the two regimes.

Summarizing this analysis of acute morbidity in adults due to all causes, several points emerge. First, regardless of how it was measured, the coefficient on the ozone variable was positive in every regression explaining minor restricted activity days in adults. It was significant at at least the 10 percent level in about half the models, although its significance did fall when the sample was restricted to self-respondents, for whom reported morbidity data were perhaps the most accurate. No evidence was found linking ozone to work loss or bed disability. Total suspended particulates were also always positively and sometimes significantly related to minor restrictions in activity.[8] Sulfates were consistently negative and significantly associated with MRADs, a counter-intuitive finding. Of all the regressors, the dummy variable indicating a chronic illness was by far the most significant.

4.1.2 Children

Children are of particular importance because they are thought to be more sensitive than adults to the effects of ozone and other air pollutants.[9] This is not only because their bodies are developing structurally, functionally, and biochemically, but also because they tend to be outdoors more than the general population, exercise more heavily, and ventilate more per unit body weight than adults. Using subsets of the 15,711 children in the final RFF data base, a number of models were explored linking varying types of acute morbidity to air pollution and other determinants.

The results of the analysis of children's MRADs due to all causes during the reference period are presented in Table 4-4. The number of explanatory variables used in the analysis is smaller because height and weight were not recorded for children in the HIS (thus necessitating exclusion of FAT and FATSQ), nor is the marital status or smoking dummy relevant (because no one under the age of 17 was given the smoking supplement).

Equation (28) includes ozone and sulfates as the relevant measures of ambient air pollution. Although both had a negative and therefore unexpected sign, neither was significant. Boys had more MRADs during the reference period than girls, a finding significant at well above the 5 percent level. The coefficients on AGE and AGESQ suggest that MRADs among children decrease with age from birth to about ten years and then increase between age ten and age sixteen, the point at which youths become adults in the HIS framework. Once again the chronic limitation dummy was the most significant determinant of acute morbidity as measured by MRADs.

Equations (29) and (30) reproduce this basic model using different combinations of air pollutants. None of the pollutants was significantly related to MRADs in children. Equations (31) - (33) reiterate this approach using the square of the average daily maximum 1-hour concentration as the measure of ozone. Unlike the adult regressions, use of the squared term had little effect. Over the six different models estimated in (28) - (33), only nitrogen dioxide showed a positive association with MRADs in children, a finding not significant at the 10 percent level.

Table 4-4. Restricted Activity Among Children: All Causes

Model #	(28)	(29)	(30)	(31)	(32)	(33)
Dep. Var. Indep. Var.	RADMINOR	RADMINOR	RADMINOR	RADMINOR	RADMINOR	RADMINOR
INTERCEPT	.165 (1.24)	.160 (1.16)	.184 (1.19)	.171 (1.28)	.168 (1.22)	.195 (1.25)
O3NR01	-.802 (1.33)	-.819 (1.35)	-.684 (1.03)	--	--	--
O3NR01SQ	--	--	--	-2.692 (0.84)	-2.736 (0.84)	-1.849 (0.55)
S4NR01	-.00284 (1.29)	-.00287 (1.28)	-.00289 (1.08)	-.00306 (1.40)	-.00306 (1.37)	-.00306 (1.14)
SPNR01	--	.0000604 (0.14)	-.000110 (0.23)	--	.0000345 (0.08)	-.000141 (0.29)
N2NR01	--	--	.000446 (1.41)	--	--	.000397 (1.27)
CONR01	--	--	-.00329 (0.64)	--	--	-.00303 (0.59)
RACEW1B0	.0178 (0.51)	.0178 (0.52)	.0107 (0.28)	.0167 (0.48)	.0167 (0.48)	.00952 (0.25)
SEXM1F0	.0633 (2.51)	.0628 (2.49)	.0497 (1.76)	.0633 (2.51)	.0629 (2.49)	.0496 (1.75)
INCOMCON	4.444E-7 (0.29)	4.438E-7 (0.29)	.00000183 (1.06)	4.812E-7 (0.32)	4.718E-7 (0.31)	.00000186 (1.08)
CITYI100	-.0161 (0.58)	-.0159 (0.57)	-.0173 (0.54)	-.0140 (0.50)	-.0137 (0.49)	-.0146 (0.46)
AGE	-.0476 (4.74)	-.0477 (4.74)	-.0496 (4.39)	-.0476 (4.74)	-.0476 (4.74)	-.0496 (4.39)
AGESQ	.00225 (3.83)	.00226 (3.83)	.00246 (3.71)	.00225 (3.82)	.00226 (3.83)	.00246 (3.71)
CHRLMDUM	.375 (5.41)	.375 (5.41)	.369 (4.71)	.376 (5.42)	.376 (5.42)	.370 (4.72)

Table 4-4 (Continued). Restricted Activity Among Children: All Causes

Model #	(28)	(29)	(30)	(31)	(32)	(33)
Dep. Var. / Indep. Var.	RADMINOR	RADMINOR	RADMINOR	RADMINOR	RADMINOR	RADMINOR
CROWDING	-.0253	-.0257	.0223	-.0262	-.0265	.0218
	(0.65)	(0.66)	(0.53)	(0.67)	(0.68)	(0.52)
AVMAXTMP	.00105	.00106	.000279	.000573	.000575	-.000176
	(1.14)	(1.14)	(0.26)	(0.71)	(0.71)	(0.19)
AVPRECIP	.111	.114	.223	.127	.129	.236
	(0.84)	(0.86)	(1.46)	(0.97)	(0.97)	(1.55)
HUMIDRF	.00149	.00150	.000193	.00143	.00143	.000177
	(0.94)	(0.94)	(0.10)	(0.90)	(0.89)	(0.10)
R^2	.0131	.0131	.0133	.0129	.0129	.0131
F OR X^2	5.58	5.16	3.44	5.50	5.08	3.39
N	5,480	5,473	4,096	5,480	5,473	4,096

Table 4-5 illustrates the results of the basic model applied to school loss among children. While the coefficient on ozone was positive equations (34) - (39) its level of significance was always very low. With the exception of nitrogen dioxide, none of the air pollutants played a significant role in explaining children's school loss. In both equation (36) and (39) nitrogen dioxide is positively associated with school loss, an effect significant at the 5 percent level.

Among the other variables, race was generally significant at at least the 10 percent level; the dummy variable indicates that white children experience more school loss due to acute morbidity than black children, holding income and other factors constant. As usual, the presence of a chronic illness has the strongest association with acute morbidity leading to school loss. Average maximum temperature was negatively and significantly related to school loss. This may reflect a reduced incidence of influenza and other flu-like diseases in warmer areas during the school year. The dummy variable for central city residence was significant at the 10 percent level in almost all these regressions, indicating more school loss days for central city students than for those outside the central city but still in an SMSA.

In equations (40) and (41) logistic regression was employed to examine the sensitivity of the school loss results to the method of ordinary least squares. Neither ozone nor other air pollutants played a significant role in explaining the variation in school loss days. As in the ordinary least squares estimates, the race of the student and the presence of a chronic

Table 4-5. School Loss Among Children: All Causes

Model #	(34)	(35)	(36)	(37)	(38)	(39)	(40)	(41)
Dep. Var. Indep. Var.	WKSLLD2W	WKSLLD2W	WKSLLD2W	WKSLLD2W	WKSLLD2W	WKSLLD2W	WSLPRES (LOGIT)	WSLPRES (LOGIT)
INTERCEPT	.542 (1.56)	.553 (1.49)	.759 (1.94)	.543 (1.56)	.522 (1.47)	.753 (1.82)	-2.733 (2.77)	-2.66 (2.75)
O3NR01	1.029 (0.75)	1.009 (0.73)	.900 (0.56)	--	--	--	-.476 (0.10)	-.343 (0.10)
O3NR01SQ	--	--	--	5.150 (0.49)	5.981 (0.48)	3.786 (0.27)	--	--
S4NR01	-.00227 (0.54)	-.00239 (0.64)	-.000103 (0.02)	-.00211 (0.60)	-.00227 (0.61)	-.0000319 (0.01)	.00248 (0.24)	.00396 (0.42)
SPNR01	--	.0000862 (0.12)	-.000366 (0.55)	--	.000112 (0.16)	-.000530 (0.51)	--	-.000914 (0.37)
N2NR01	--	--	.00115 (2.07)	--	--	.00118 (2.14)	--	--
CONR01	--	--	-.00309 (0.35)	--	--	-.00953 (1.00)	--	--
RACEW180	.115 (1.92)	.114 (1.92)	.0710 (1.01)	.116 (1.94)	.115 (1.93)	.0719 (1.02)	.628 (3.29)	.625 (3.27)
SEXM1F0	-.0497 (1.16)	-.0477 (1.13)	-.0592 (1.39)	-.0499 (1.16)	-.0479 (1.13)	-.0690 (1.38)	-.223 (1.84)	-.215 (1.77)
INCOMCON	-.00000160 (0.64)	-.00000153 (0.61)	-.00000281 (0.94)	-.00000151 (0.65)	-.00000153 (0.51)	-.00000292 (0.94)	-.00000227 (0.33)	-.00000239 (0.33)
CITYI130	.0930 (1.93)	.0930 (1.92)	.0743 (1.32)	.0875 (1.40)	.0874 (1.83)	.0717 1.28	--	--
AGE	-.0122 (0.23)	-.0101 (0.19)	-.0273 (0.43)	-.0125 (0.23)	-.0104 (0.20)	-.0276 (0.44)	.0727 (0.47)	.0878 (0.57)
AGESQ	.000595 (0.25)	.000511 (0.21)	.00151 (0.53)	.000612 (0.26)	.000525 (0.22)	.00152 (0.54)	-.00309 (0.45)	-.00370 (0.53)
CHRLMDUM	.731 (7.38)	.724 (7.23)	.729 (5.32)	.731 (7.33)	.724 (7.23)	.719 (6.01)	1.137 (5.48)	1.102 (5.26)

Reproduced from best available copy.

Table 4-5 (Continued). School Loss Among Children: All Causes

Model #	(34)	(35)	(36)	(37)	(38)	(39)	(40)	(41)
Dep. Var.								
Indep. Var.	WKSLLD2W	WKSLLD2W	WKSLLD2W	WKSLLD2W	WKSLLD2W	WKSLLD2W	WSLPRES (LOGIT)	WSLPRES (LOGIT)
CROWDING	-.0330 (0.48)	-.0332 (0.48)	-.0399 (1.24)	-.0309 (0.45)	-.0312 (0.45)	-.0976 (1.21)	--	--
AVMAXTMP	-.00480 (2.77)	-.00483 (2.76)	-.00479 (2.28)	-.00441 (2.75)	-.00446 (2.74)	-.00435 (2.26)	-.00911 (1.87)	-.00874 (1.77)
AVPRECIP	-.00356 (0.02)	.00352 (0.02)	.122 (0.45)	-.00915 (0.04)	-.000137 (0.00)	.144 (0.43)	.483 (0.81)	.425 (0.63)
HUMIDRF	-.000546 (0.20)	-.000494 (0.19)	-.00250 (0.77)	-.000499 (0.19)	-.000434 (0.16)	-.00248 (0.76)	.00358 (0.46)	.00321 (0.41)
R2	.0243	.0237	.0255	.0241	.0235	.0254		
F or X2	5.47	4.96	3.41	5.44	4.93	3.39	50.17	48.58
N	2,874	2,870	2,999	2,874	2,870	2,038	2,888	2,884

limitation were the most significant factors, with average maximum temperature significant at the 10 percent level.

Concluding the analysis of acute morbidity in children due to all causes, Table 4-6 presents the results of the analysis of children's bed disability during the reference period. Equations (42) - (44) utilize several combinations of other pollutants entered with ozone in its linear form. In (45) - (47), the square of ozone is utilized. Although ozone was positive in all six regressions, it was never significant. Nor were any of the other air pollution variables, including nitrogen dioxide which was found to be significantly associated with school loss days--a less serious form of acute morbidity than a bed disability day. In the analysis of bed disability, age and its square were both significant, as was the presence of a chronic limitation. Average maximum temperature during the reference period was also negative and significant, as in the school loss regressions. This again may be due to the relatively lower likelihood of influenza in warmer areas. Finally, the central city residence dummy was again a significant determinant of bed disability, almost always at the 5 percent level.

4.2 Acute Morbidity Due to Respiratory Disease

Respiratory disease is that type most likely to result from exposure to ozone and other air pollutants in ambient environments. Before undertaking the analysis of acute respiratory morbidity in adults and children, an additional requirement was imposed on the air pollution data. Besides requiring that any ozone or other air pollution data be derived

Table 4-6. Bed Disability Among Children: All Causes

Model #	(42)	(43)	(44)	(45)	(46)	(47)
Dep. Var. Indep. Var.	BEDDIS2W	BEDDIS2W	BEDD9S2W	BEDDIS2W	BEDDIS2W	BEDDIS2W
INTERCEPT	.826 (6.42)	.796 (5.97)	.873 (5.39)	.820 (6.36)	.787 (5.88)	.858 (5.29)
O3NR01	.874 (1.49)	.848 (1.43)	.931 (1.33)	--	--	--
O3NR01SQ	--	--	--	3.223 (1.03)	2.913 (0.92)	2.604 (0.74)
S4NR01	-.00214 (1.02)	-.00246 (1.14)	-.00343 (1.22)	-.00192 (0.92)	-.00226 (1.05)	-.00320 (1.14)
SPNR01	--	.000261 (.062	.0000150 (0.03)	--	.00286 (0.68)	.000555 (0.11)
N2NR01	--	--	.000450 (1.36)	--	--	.000513 (1.56)
CONR01	--	--	-.00273 (0.51)	--	--	-.00310 (.058)
RACEW1B0	.0481 (1.45)	.0476 (1.41)	.0613 (1.52)	.0493 (1.46)	.0487 (1.45)	.0627 (1.56)
SEXM1F0	-.000335 (0.01)	.000910 (0.04)	.0133 (0.45)	-.000387 (0.02)	.000840 (0.03)	.0135 (0.46)
INCOMCON	-2.024E-7 (0.14)	-2.756E-7 (0.19)	6.736E-7 (0.04)	-2.470E-7 (0.17)	-3.113E-7 (0.21)	2.089E-8 (0.01)
CITYI100	.0550 (2.05)	.0527 (1.96)	.0754 (2.26)	.0528 (1.97)	.0503 (1.88)	.0716 (2.15)
AGE	-.0304 (3.14)	-.0307 (3.17)	-.0430 (3.65)	-.0304 (3.14)	-.0308 (3.17)	-.0430 (3.65)
AGESQ	.00147 (2.59)	.00150 (2.64)	.00215 (3.11)	.00148 (2.60)	.00150 (2.64)	.00215 (3.10)
CHRLMDUM	.486 (7.81)	.489 (7.84)	.603 (7.86)	.485 (7.79	.488 (7.83)	.601 (7.85)

Table 4-6 (Continued). Bed Disability Among Children: All Causes

Model #	(42)	(43)	(44)	(45)	(46)	(47)
Dep. Var. Indep. Var.	BEDDIS2W	BEDDIS2W	BEDD9S2W	BEDDIS2W	BEDDIS2W	BEDDIS2W
CROWDING	-.0142 (0.37)	-.0167 (0.44)	-.0423 (0.94)	-.0131 (0.34)	-.0157 (0.41)	-.0414 (0.92)
AVMAXTMP	-.00513 (5.74)	-.00520 (5.81)	-.00592 (5.39)	-.00463 (5.99)	-.00471 (6.06)	-.00531 (5.58)
AVPRECIP	.233 (1.85)	.246 (1.92)	.313 (1.97)	.219 (1.74)	.232 (1.81)	.296 (1.87)
HUMIDRF	-.00506 (3.31)	-.00466 (3.00	-.00495 (2.54)	-.00500 (3.26)	-.00457 (2.94)	-.00491 (2.51)
R^2	.0197	.0199	.0271	.0196	.0197	.0269
F or X^2	9.30	8.70	7.81	9.21	8.61	7.73
N	6.018	6.008	4.495	6.018	6.008	4.495

from monitors no further than twenty miles from an individual's home, a requirement was added concerning the completeness of the ozone data during the two-week reference period. Approximately 400 individuals were excluded from the following analyses because they were matched to an ozone monitor which reported fewer than 168 of the 336 possible hourly readings during the two-week period (24 hours x 14 days). The following results, then, are predicated on the air pollution monitors nearest the individuals' houses, all within twenty miles, as well as on thorough reporting at each monitor.

4.2.1 Adults

The first results to be presented concern the total number of restricted activity days (TRADs) due to respiratory disease during the reference period. This variable differs somewhat from the MRADs variable used in the "all causes" analysis. There attention was centered on individuals who reported RADs but no work loss or bed disability days so we could be sure RADs were "minor." In the present analysis RADs may include work or bed disability due to respiratory illness. This alternative approach was adopted because relatively few individuals reported a respiratory illness that did not result in work loss or bed disability (both of which are analyzed separately below).

Table 4-7 presents the analysis of total restricted activity days in adults due to respiratory illness. The general approach is that adopted in the "all causes" analysis. First, simple ordinary least squares was used with several different combinations of air pollution variables. Then variants of this plus other approaches were employed.

Table 4-7. Restricted Activity Days Among Adults: Respiratory Disease

Model #	(48)	(49)	(50)	(51)	(52)	(53)
Dep. Var. Indep. Var.	TRADRSP	TRADRSP	TRADRSP	TRADRSP	RADRSPPS (LOGIT)	RADRSPPS (LOGIT)
INTERCEPT	.724 (1.95)	.536 (1.44)	.791 (1.75)	.561 (1.51)	-3.939 (2.61)	-3.972 (2.41)
O3NR01	1.544 (1.98)	1.873 (2.32)	1.409 (1.46)	--	6.231 (1.79)	3.650 (0.92)
O3NR01SQ	--	--	--	7.479 (1.75)	--	--
S4NR01	-.00334 (1.08)	-.00154 (0.48)	-.00255 (0.61)	-.00128 (0.40)	.000578 (0.85)	.000525 (0.76)
SPNR01	--	-.000302 (0.53)	-.000594 (0.84)	-.000235 (0.41)	-.00234 (0.00)	-.00241 (0.00)
N2NR01	--	--	.000671 (1.48)	--	--	.00172 (0.96)
CONR01	--	--	.00476 (0.64)	--	--	.0168 (0.54)
RACEW1B0	.103 (1.99)	.100 (1.94)	.108 (1.75)	.103 (2.00)	.454 (1.77)	.379 (1.37)
SEXM1F0	.0403 (1.10)	.0384 (1.04)	.0430 (0.96)	.0389 (1.06)	-.126 (0.75)	-.106 (0.57)
MARY1N0	-.0177 (0.47)	-.0186 (0.49)	.0167 (0.36)	-.0165 (0.43)	-.352 (2.10)	-.247 (1.31)
INCOMCON	-.00000470 (2.37)	-.00000465 (2.34)	-.00000627 (2.62)	-.00000465 (2.34)	-.0000126 (1.40)	-.0000239 (2.28)
FAT	-.605 (2.32)	-.601 (2.30)	-.795 (2.51)	-.600 (2.30)	-.951 (0.99)	-1.272 (1.26)
FATSQ	.120 (2.39)	.119 (2.37)	.161 (2.63)	.119 (2.37)	.175 (0.97)	.239 (1.29)
AGE	.00981 (1.77)	.0101 (1.82)	.0105 (1.58)	.00996 (1.80)	.0408 (1.55)	.0601 (2.02)

Table 4-7 (continued) Restricted Activity Days Among Adults: Respiratory Disease

Model #	(48)	(49)	(50)	(51)	(52)	(53)
Dep. Var. Indep. Var.	TRADRSP	TRADRSP	TRADRSP	TRADRSP	RADRSPPS (LOGIT)	RADRSPPS (LOGIT)
AGESQ	-.000111 (1.89)	-.000113 (1.94)	-.000112 (1.59)	-.000112 (1.92)	-.000553 (1.92)	-.000701 (2.19)
SMOKY1NO	.0371 (1.07)	.0364 (1.04)	.0466 (1.10)	.0361 (1.03)	.255 (1.67)	.219 (1.25)
EDCOMCON	.00206 (0.36)	.00177 (0.31)	-.00121 (0.18)	.00175 (0.31)	.00903 (0.35)	.0130 (0.45)
CHRLMDUM	.247 (5.29)	.247 (5.29)	.256 (4.51)	.245 (5.25)	.565 (2.98)	.457 (2.14)
AVMAXTMP	-.00401 (3.26)	--	--	--	--	--
DMINTEMP	--	-.00138 (0.83)	-.00233 (1.14)	-.000669 (0.42)	-.00514 (0.79)	-.00765 (1.01)
DMAXTEMP	--	-.00265 (2.67)	-.00316 (2.61)	-.00193 (2.16)	-.0146 (3.38)	-.0142 (2.91)
AVPRECIP	.119 (0.70)	.128 (0.73)	-.000492 (0.00)	.104 (0.60)	.275 (0.36)	-.0914 (0.10)
HUMIDRF	.00190 (0.91)	.00243 (1.17)	.00228 (0.86)	.00231 (1.11)	.0250 (2.50)	.0264 (2.28)
R^2	.0142	.0150	.0190	.0146		
F or X^2	4.41	4.14	3.56	4.00	73.88	57.11
N	4,906	4,899	3,703	4,899	4,899	3,703

In equations (48) - (50), ozone was positively related to the number of RADs due to respiratory disease. In the first two regressions this effect was significant at the 5 percent level; in the third, where five pollutants are included, the coefficient was insignificant. Income was negatively and significantly related to respiratory RADs. So was average maximum temperature in equation (48) and DMAXTEMP in equations (49) and (50). (The latter measures average daily maximum temperature if the respondent was interviewed during the second and third quarters of the year, and equals zero if the respondent was interviewed during the first or fourth quarters. This distinction was introduced because higher temperatures in the winter may be good while those in the summer may be harmful.) Both FAT and FATSQ entered the equation significantly, suggesting that there is some ideal weight-to-height ratio which minimizes respiratory RADs. Whites had more respiratory RADs than blacks, a difference that was quite significant. The chronic limitation dummy once again was significantly associated with acute morbidity as measured by respiratory RADs.

As in the "all causes" analysis, we experimented with a squared ozone term. In general, it did not outperform the simple linear measure. Equation (51) represents an illustrative case. The square of ozone was significant at the 10 percent level as contrasted with the otherwise identical equation (49) where the linear ozone term was significant at above the 5 percent level. Equations (52) and (53) are logit analyses of the likelihood that an individual will have at least one respiratory RAD. They are consistent with the OLS results although the ozone variable was somewhat less significant in the logit analysis than in OLS. As expected, most variables exhibit similar effects, although FAT and FATSQ were no longer significant in the logistic framework.

The performance of the dummy variable for smoking is worth mentioning here. Even though the focus in equations (48) - (53) was on respiratory disease, the smoking dummy was generally less significant than it was in the all-causes regressions. Although its sign was always positive, only in equation (52) was it significant and even there only at the 10 percent level. One possible explanation for this may be that individuals with chronic respiratory illnesses, who would account for a substantial share of acute respiratory illness, have quit smoking on their own or on a doctor's instructions. Thus, they would not have been identified as current smokers in the HIS even though they once smoked. In subsequent sensitivity analysis, we probe more deeply into this possibility.

The results from the analysis of work loss due to respiratory disease are presented in Table 4-8. Equations (54) and (55) are OLS estimates using ozone with two combinations of other pollutants. None of the air pollutants was significant in either regression with the exception of particulates (at the 10 percent level) which had an unexpected sign. In (56) and (57) the square of the average daily maximum ozone concentration was used. This substitution changed the results little, if at all. Among the other explanatory variables only education was significant at the 5 percent level--the sign indicating that individuals with more education have fewer work loss days due to respiratory disease. Even the presence of

Table 4-8. Adult Work Loss: Respiratory Disease

Model #	(54)	(55)	(56)	(57)	(58)	(59)	(60)	(61)
Dep. Var.								
Indep. Var.	TWSLRSP	TWSLRSP	TWSLRSP	TWSLPSP	TWSLRSP	TWSLRSP	WSLRSPPS (LOGIT)	WSLRSPPS (LOGIT)
INTERCEPT	.123 (0.52)	.0496 90.16)	.117 (0.50)	.0391 (0.13)	.197 (0.92)	1.05 (0.35)	-2.925 (1.18)	-3.149 (1.13)
D3NR01	.0713 (0.16)	-.0179 (0.03)	--	--	.103 (0.24)	.0277 (0.50)	-1.081 (0.17)	-1.627 (0.22)
D3NR01SQ	--	--	-.498 (0.21)	-1.236 (0.45)	--	--	--	--
S4NR01	-.00112 (0.57)	-.00236 (0.86)	-.00111 (0.57)	-.00234 (0.84)	-.000624 (0.33)	-.00169 (0.65)	-.0169 (0.62)	-.0414 (1.15)
SPNR01	-.000457 (1.29)	-.000802 (1.72)	-.000427 (1.21)	-.000756 (1.63)	-.000479 (1.40)	-.000834 (1.85)	-.0000868 (0.00)	-.00180 (0.32)
N2NR01	--	.000357 (1.27)	--	.000336 (1.41)	--	.000383 (1.40)	--	.00391 (1.32)
CONR01	--	.00342 (0.72)	--	.00322 (0.68)	--	.00324 (0.70)	--	.0246 (0.43)
RACEW1B0	.0147 (0.47)	.00977 (0.25)	.0162 (0.49)	.0194 (0.25)	.00352 (0.12)	-.00237 (0.09)	.310 (0.69)	.0959 (0.20)
SEXM1F0	-.00955 (0.41)	-.0193 (0.61)	-.00933 (0.40)	-.0181 (0.61)	-.00405 (0.18)	-.0114 (0.40)	-.0485 (0.17)	-.114 (0.35)
MARYIN0	-.0169 (0.73)	-.0131 (0.44)	-.0167 (0.72)	-.0129 (0.43)	-.00680 (0.31)	-.00143 (0.05)	-.195 (0.64)	-.124 (0.37)
INCOMCON	-9.681E-7 (0.90)	-.00000126 (0.82)	-9.601E-7 (0.79)	-.00000125 (0.80)	-9.320E-7 (0.80)	-.00000129 (0.85)	-.00000237 (0.14)	-.00000592 (0.35)
FAT	-.0223 (0.14)	.0416 (0.23)	-.0219 (0.13)	.0436 (0.23)	-.0383 (0.53)	-.0379 (0.19)	-1.899 (1.39)	-1.737 (1.11)
FATSQ	.00319 (0.10)	-.00832 (0.22)	.00301 (0.10)	-.00822 (0.22)	.0162 (0.34)	.00503 (0.13)	.350 (1.49)	.246 (1.23)
AGE	.00419 (0.78)	.00431 (0.73)	.00413 (0.90)	.00432 (0.78)	.00387 (0.70)	.00295 (0.55)	.0101 (0.17)	.00221 (0.00)

Table 4-8 (Continued). Adult Work Loss: Respiratory Disease

Model #	(54)	(55)	(56)	(57)	(58)	(59)	(60)	(61)
Dep. Var. Indep. Var.	TWSLRSP	TWSLRSP	TWSLRSP	TWSLRSP	TWSLRSP	TWSLRSP	WSLRSPPS (LOGIT)	WSLRSPPS (LOGIT)
AGESQ	-.0000553	-.0000575	-.0000555	-.0000577	-.0000406	-.0000410	-.000252	-.0000708
	(1.11)	(0.90)	(1.11)	(0.50)	(0.85)	(0.67)	(0.36)	(0.10)
SMOKYIND	-.0263	-.0172	-.0264	-.0171	-.0205	-.00934	-.124	.127
	(1.30)	(0.65)	(1.30)	(0.65)	(1.060	(0.37)	(0.48)	(0.44)
EDCOMCGN	-.00943	-.0106	-.00944	-.0106	-.00296	-.00155	.00229	.0504
	(2.65)	(2.30)	(2.65)	(2.31)	(0.79)	(0.35)	(0.00)	(0.87)
CHRLNDUM	-.000196	-.00731	-.000164	-.00730	.0122	.00995	.245	.247
	(0.01)	(0.17)	(0.01)	(0.17)	(0.37)	(0.23)	(0.58)	(0.54)
DMINTEMP	.000198	-.0000279	.000263	-.00000669	.000254	.0000719	-.00509	-.0106
	(0.20)	(0.02)	(0.28)	(0.01)	(0.27)	(0.05)	(0.48)	(0.82)
DMAXTEMP	-.000530	-.000548	-.000452	-.000499	-.000477	-.000474	-.0130	-.0122
	(0.91)	(0.72)	(0.86)	(0.71)	(0.85)	(0.65)	(1.79)	(1.49)
AVPRECIP	-.0403	-.0170	-.0471	-.0255	-.0123	.0251	-.372	.678
	(0.39(	(0.13)	(0.46)	(0.19)	(0.13)	(0.19)	(0.28)	(0.47)
HUMIDRF	.00189	.00184	.00195	.00193	.00113	.000704	.0352	.0241
	(1.54)	(1.10)	(1.58)	(1.15)	(0.96)	(0.44)	(2.13)	(1.28)
BLUE	--	--	--	--	.0327	.0516	.221	.461
					(1.51)	(1.86)	(0.77)	(1.44)
R2	.0089	.0106	.0063	.0107	.0072	.0034		
F or X2	1.46	1.17	1.46	1.18	1.09	0.97		
N	2,943	2,205	2,343	2,205	2,374	2,152	2,874	2,152

Reproduced from best available copy.

a chronic illness or limitation does not play a significant role in explaining respiratory work loss.

Recalling the effect that a blue collar dummy had in the all-causes work loss equations, several models were estimated testing its effects on respiratory work loss. Two of these regressions are presented in (58) and (59) in Table 4-8. It had the anticipated sign in both equations and was significant at the 10 percent level in (59). Again, this suggests that blue collar jobs might have health or accident risks that affect those working at them. In these models, education was no longer significant, possibly due to collinearity between education and occupation. Finally, the sensitivity of these OLS results was explored using logistic regression. Equations (60) and (61) present the results of logit models which include the blue collar dummy variable. The analysis indicates no significant effect of any of the air pollution variables.

Before leaving the analysis of work loss due to acute respiratory disease, it is important to comment on the F-statistics of the regressions presented. They were low. In none of the regressions can one accept at the 5 percent level the hypothesis that the entire set of regressors (with the exception of the constant term) has any effect in explaining variation in work loss due to respiratory causes. These are the only results of all those analyzed for acute illness where this is the case. The problem may be due in part to the relative infrequency of respiratory related work loss in our sample. The average number of respiratory-related work loss days during the two-week reference period was 0.05 (in other words, one respiratory work loss day every forty weeks). In the logistic analysis, of the 3,000 or so individuals included in the different runs, only 50 to 60 had at least one respiratory work loss day during the reference period.

Table 4-9 shifts attention to the analysis of bed disability days due to acute respiratory disease. In keeping with the usual custom the first three regressions, equations (62) - (64), present ordinary least squares estimates in which the average daily maximum ozone concentration over the two week period was entered linearly in the company of different combinations of air pollutants and other explanatory variables. A fourth equation, number (65), used the square of ozone to explore the kinds of non-linear effects suggested by the earlier analysis of restricted activity days due to all causes. While always positive, ozone was not significant in these four bed disability regressions. Neither were any of the other air pollutants with the exception of nitrogen dioxide in equation (64) where it was positively associated with bed disability and significant at the 10 percent level.

Among the non-pollution variables, income, general physical condition (FAT), maximum temperature (in the summer), and the chronic limitation dummy were also significant--as they have been in many of the other models investigated to this point. Age and its square moved in and out of significance depending on the particular specification of the model. Note that the present smoking status of those in the RFF sample, while always positive, was never significant.

Equations (66) - (69) contain the results of logistic analysis of adult bed disability due to respiratory disease, using both ozone and its square. The results did not differ much from those obtained using ordinary

Table 4-9. Adult Bed Disability: Respiratory Disease

Model #	(62)	(63)	(64)	(65)	(66)	(67)	(68)	(69)
Dep. Var. / Indep. Var.	TBEDRSP	TBEDRSP	TBEDRSP	TBEDRSP	BEDRSPPS (LOGIT)	BEDRSPPS (LOGIT)	BEDRSPPS (LOGIT)	BEDRSPPS (LOGIT)
INTERCEPT	.630 (3.99)	.823 (3.76)	1.139 (4.18)	.822 (3.75)	-4.234 (2.45)	-3.432 (1.30)	-4.196 (2.39)	-3.377 (1.77)
O3NR01	.0125 (0.04)	.404 (0.85)	.0474 (0.08)	--	4.255 (0.950	1.643 (0.32)	--	--
O3NR01SQ	--	--	--	.746 (0.30)	--	--	20.143 (0.95)	10.763 (0.46)
S4NR01	--	-.000100 (0.05)	-.00000762 (0.00)	-.0003428 (0.02)	-.00652 (0.37)	-.0181 (0.77)	-.00615 (0.35)	-.0179 (0.77)
SPNR01	.000233 (1.07)	.000138 (0.41)	-.000113 (0.28)	.000175 (0.52)	.00285 (0.27)	.000891 (0.24)	.00292 (0.92)	.000840 (0.22)
N2NR01	--	--	.000479 (1.76)	--	--	.00321 (1.55)	--	.00332 (1.58)
CONR01	--	--	.00392 (0.97)	--	--	.0304 (0.93)	--	.0305 (0.84)
RACEW130	.0219 (0.99)	.0267 (0.89)	.0352 (0.35)	.0278 (0.92)	.150 (0.55)	.0497 (0.17)	.155 (0.57)	.0492 (0.17)
SEXM1FO	-.00265 (0.17)	.0313 (1.44)	.0451 (1.58)	.0315 (1.45)	-.291 (1.37)	-.212 (0.90)	-.291 (1.37)	-.212 (0.90)
MARYIND	-.00309 (0.45)	-.0304 (1.36)	-.0165 (0.50)	-.0208 (1.32)	-.541 (2.59)	-.403 (1.72)	-.534 (2.56)	-.402 (1.72)
INCOMCON	-.00000260 (3.060	-.00000311 (2.66)	-.00000337 (2.34)	-.00000310 (2.65)	-.0000179 (1.65)	-.0000251 (1.71)	-.0000190 (1.66)	-.0000252 (1.71)
FAT	-.513 (4.39)	-.752 (4.83)	-.1025 (5.39)	-.752 (4.83)	-1.125 (1.10)	-1.454 (1.53)	-1.128 (1.11)	-1.457 (1.53)
FATSQ	.103 (4.57)	.152 (5.13)	.210 (5.71)	.152 (5.12)	.231 (1.24)	.241 (1.79)	.231 (1.27)	.342 (1.79)
AGE	.00220 (3.32)	.00475 (2.07)	.00704 (1.75)	.00674 (2.04)	.0652 (1.37)	.0799 (2.02)	.0552 (1.94)	.0795 (2.01)

Reproduced from best available copy.

Table 4-9 (continued)

Model #	(62)	(63)	(64)	(65)	(66)	(67)	(68)	(69)
Dep. Var. / Indep. Var.	TBEDRSP	TBEDRSP	TBEDRSP	TBEDRSP	BEDRSPPS (LOGIT)	BEDRSPPS (LOGIT)	BEDRSPPS (LOGIT)	BEDRSPPS (LOGIT)
AGESQ	-.0000288 (1.14)	-.0000777 (2.25)	-.0000793 (1.89)	-.0000776 (2.25)	-.000994 (2.44)	-.00109 (2.46)	-.000985 (2.44)	-.00109 (2.45)
SMOKYIND	.00335 (0.52)	.0127 (0.62)	.0174 (0.58)	.0126 (0.61)	.0689 (0.35)	.0665 (0.30)	.0692 (0.36)	.0665 (0.32)
EDCOMCON	-.0000500 (0.02)	-.000931 (0.23)	-.00223 (0.54)	-.000931 (0.29)	-.0235 (0.71)	-.0241 (0.67)	-.0236 (0.71)	-.0240 (0.66)
CHRLMDUM	.0771 (3.88)	.0783 (2.84)	.0331 (2.52)	.0779 (2.83)	.432 (1.74)	.369 (1.35)	.430 (1.75)	.367 (1.35)
DMINTEMP	-.000669 (2.39)	-.000771 (2.39)	-.00101 (1.93)	-.000578 (2.27)	-.0135 (3.75)	-.0194 (3.31)	-.0121 (3.97)	-.0189 (3.56)
AVPRECIP	.0353 (0.50)	-.00706 (0.07)	-.00427 (0.35)	-.0193 (0.18)	.154 (0.14)	.505 (0.46)	.122 (0.14)	.516 (0.47)
HUMIDRF	.00119 (1.43)	.00164 (1.34)	.000991 (0.62)	.00167 (1.36)	.0396 (3.13)	.0319 (2.24)	.0393 (3.08)	.0314 (2.20)
R2	.0097	.0156	.0202	.0154				
F or X2	4.84	4.29	3.80	4.25	85.05	64.20	85.97	64.28
N	4,382	4,839	3,703	4,839	4,839	3,703	4,899	3,703

least squares. None of the air pollution variables was significant at the 5 percent level in the logit specifications. The dummy variable for marital status was significant in the logit specification although it was not in OLS. It indicates that married persons are less likely than singles to experience respiratory bed disability days. The other results were similar to those obtained in OLS as well.

In reviewing the analysis of acute morbidity among adults due to respiratory disease, one conclusion emerges. Although there appears to be no connection between average daily maximum 1-hour concentrations of ozone and either work loss or bed disability, the results suggest that there may be a positive and statistically significant relationship between ozone and total restricted activity days due to respiratory disease. In two of the regressions in Table 4-7 the coefficient on ozone was significant at the 5 percent level and in two other cases it was significant at the 10 percent level. It is important to examine this association in more detail.

A considerable number of sensitivity analyses were performed on the association between the average daily maximum 1-hour concentration of ozone and acute respiratory morbidity in the form of restricted activity days. A number of these results are reported in Table 4-10.

To begin with, we considered the possibility that imposing a requirement for at least 168 hours of valid ozone data during the reference period may have biased the results (this might be the case, for instance, if the monitors in the highest ozone areas were the ones most likely to have complete data). In equation (70) we duplicated the model in (49) (Table 4-7) but included the 400 individuals whose ozone data did not meet

Table 4-10. Sensitivity Analysis of Adult Restricted Activity: Respiratory Disease

Model #	(70)	(71)	(72)	(73)	(74)	(75)	(76)	(77)
Dep. Var. / Indep. Var.	TRADRSP	TRADRSP	TRADRSP	TRADRSP	RADRSPPS (LOGIT)	TRADRSP	TRADRSP	TRADRSP
INTERCEPT	.441 (1.21)	1.159 (1.76)	.541 (1.15)	2.651 (2.63)	-1.736 (0.57)	.507 (1.29)	-.223 (0.49)	.110 (0.18)
O3NR01	1.923 (2.34)	5.370 (2.33)	1.973 (2.41)	.959 (0.83)	13.431 (2.47)	1.391 (2.31)	-- --	-- --
O310NR01	-- --	-- --	-- --	-- --	-- --	-- --	1.029 (0.90)	-- --
O305NR01	-- --	-- --	-- --	-- --	-- --	-- --	-- --	-.0763 (0.04)
S4NR01	-.00176 (0.57)	-.000235 (0.04)	-.00148 (0.39)	-.00312 (0.62)	-.0398 (1.11)	-.00101 (0.30)	-- --	-- --
S410NR01	-- --	-- --	-- --	-- --	-- --	-- --	.0000118 (0.00)	-- --
S405NR01	-- --	-- --	-- --	-- --	-- --	-- --	-- --	-0.0109 (1.76)
SPNR01	-.000295 (0.52)	.000533 (0.54)	-.00107 (1.32)	.0000613 (0.07)	-.00139 (0.25)	-.000289 (0.48)	-- --	-- --
SP10NR01	-- --	-- --	-- --	-- --	-- --	-- --	-.000368 (0.51)	-- --
SP05NR01	-- --	-- --	-- --	-- --	-- --	-- --	-- --	.000252 (0.25)
RACEW130	.104 (2.12)	.149 (1.69)	.0531 (.097)	.103 (1.33)	.934 (1.65)	.111 (2.13)	.155 (2.63)	.0645 (0.67)
SEXM1FO	.0257 (0.72)	.00832 (0.11)	.0704 (1.60)	.151 (2.43)	-.0167 (0.00)	.0404 (1.07)	.0406 (0.89)	.0729 (1.06)
MARY1NO	-.00317 (0.09)	.0374 (0.57)	-.0732 (1.55)	.107 (1.71)	-.254 (0.74)	-.0163 (0.42)	-.0231 (0.47)	-.0277 (0.39)
INCOMCON	-.00000462 (2.40)	-.00000596 (1.77)	-.00000339 (1.53)	-.0000105 (3.11)	-.0000352 (1.67)	-.00000571 (2.83)	-.00000435 (2.59)	-.00000349 (0.91)

Reproduced from best available copy.

Table 4-10 (continued)

Model #	(70)	(71)	(72)	(73)	(74)	(75)	(76)	(77)
Dep. Var. Indep. Var.	TRADRSP	TRADRSP	TRADRSP	TBEDRSP	RADRSPPS (LOGIT)	TRADRSP	TRADRSP	TRADRSP
FAT	-.524 (2.04)	-1.350 (2.97)	-.133 (0.45)	-1.527 (3.40)	-1.547 (0.94)	-.654 (2.42)	.0613 (0.20)	.00885 (0.02)
FATSQ	.104 (2.11)	.289 (3.18)	.0170 (0.30)	.295 (3.41)	.356 (1.13)	.133 (2.55)	-.0246 (0.41)	-.0175 (0.23)
AGE	.00363 (1.79)	.0114 (1.20)	.0104 (1.59)	.0117 (1.32)	.0496 (0.91)	.00395 (1.57)	.0140 (2.07)	-.00130 (0.13)
AGESQ	-.000107 (1.38)	-.000117 (1.17)	-.000119 (1.72)	-.000117 (1.25)	-.000532 (1.01)	-.000100 (1.66)	-.000155 (2.16)	.0000128 (0.12)
SMOKYIND	.0337 (1.00)	.0718 (1.121)	.0187 (0.44)	.0893 (1.52)	.0923 (0.28)	.0412 (1.15)	.0190 (0.44)	-.0100 (0.15)
EDCOMCON	.00173 (0.31)	-.00651 (0.68)	.0103 (1.47)	.00301 (0.32)	.0466 (0.93)	.000857 (0.15)	-.00342 (0.48)	.00556 (0.55)
CHRLMDUM	.252 (5.36)	.230 (2.99)	.254 (4.44)	.360 (4.58)	-- --	.251 (5.22)	.176 (3.00)	.120 (1.39)
DMINTEMP	-.0140 (0.46)	-- --	-- --	-- --	-- --	-.00256 (1.48)	-.00143 (0.68)	-.00127 (0.42)
DMAXTEMP	-.00253 (2.63)	-- --	-- --	-- --	-- --	-.00324 (3.13)	-.00196 (1.53)	-.000703 (0.37)
AVMAXTMP	-- --	-.00233 (1.02)	-.00593 (2.32)	-.00392 (1.79)	-.0278 (2.80)	-- --	-- --	-- --
AVPRECIP	.117 (0.59)	-.0669 (0.22)	.423 (2.31)	.194 (0.58)	1.079 (0.69)	.0390 (0.49)	.223 (0.98)	.590 (1.35)
HUMIDRF	.002 (1.21)	.--555 (1.60)	-.00249 (0.93)	.00398 (1.15)	.00839 (0.40)	.00265 (1.20)	.00238 (0.93)	.00203 (0.57)
POLLENRF	-- --	-- --	-- --	.00658 (0.63)	-- --	-- --	-- --	-- --
PSIYIND	-- --	-- --	-- --	-- --	-- --	.149 (1.93)	-- --	-- --
R2	.0144	.0203	.0182	.0321		.0185	.019	.0125
F or X2	4.25	2.76	2.70	4.22	22.67	4.44	2.12	0.79
N	5,264	2,284	2,497	2.314	866	4,495	3,179	1.143

the 168 hour data completeness criterion. The results were virtually identical between the two regressions.

Because the period between April and October each year is the time when ozone readings are typically high, a number of runs were performed dividing the sample in half. The purpose was to determine whether it was summer exposures alone that were responsible for the observed relationship between ozone and respiratory RADs.[10] Separate models were estimated for those interviewed in the first and fourth quarters of 1979, and those interviewed in the second and third quarters. These results are presented in equations (71) and (72), respectively, where ozone was entered with sulfates and total suspended particulates. In spite of the fact that other coefficients varied substantially between the two subsamples (particulates, the marriage dummy, and FAT and FATSQ, for instance), the coefficient on ozone was virtually identical between the two. As in equation (49), where the two subsamples were pooled, ozone was positively and significantly associated with respiratory RADs at the 5 percent level in both subsamples.

Next, we examined the sensitivity of the apparent association between RADs and ozone to the inclusion of a measure of pollen concentration (POLLENRF). In equation (73) the relationship estimated in (49) (in Table 4-7) is reproduced for individuals for whom pollen data were available. Including pollen in the regression reduced the magnitude of the coefficient on ozone by about half and reduced its t-ratio even more. This suggests that effects apparently attributable to ozone (or other air pollutants) may actually be due to allergic reactions to ragweed and other pollen.

Because this would be an important conclusion, it was pursued further. Note that the "without-pollen" respiratory RAD regression in equation (49) is based on nearly 4,900 observations. The "with-pollen" regression presented in (73) is based on 2,313 observations, less than half as many. (This difference arises because pollen data are only available for a subset of SMSAs.) This suggests the possibility that the change in the ozone coefficient may have been due to different samples rather than to a spurious correlation when pollen was omitted from the estimating equation.

To control for the heterogeneity of the samples, a simple procedure was employed. Using the same 2,313 individuals upon whom the results in (73) were based, another regression was run using respiratory RADs as the dependent variable but excluding pollen as a regressor. The results of that test (not reported here) lend credence to the view that it was the change in sample size rather than the inclusion of pollen that reduced the size and significance of the ozone coefficient. In fact, when pollen data were added to the respiratory RAD regression on the smaller sample, the coefficient on ozone increased slightly in both size and significance. Thus, the estimated coefficient on ozone may be more robust with respect to pollen concentrations than the results in equation (73) would suggest.

Besides children, another group potentially more sensitive to ozone and other air pollutants are those with chronic illnesses, particularly chronic respiratory disease.[11] For this reason, 866 individuals suffering from any type of chronic illness were identified for additional analysis. RADs due to respiratory disease among this subsample were analyzed using logit as in equation (52) above. The results of this analysis are presented in equation (74). The coefficient on ozone was more than twice as large in the "sensitive population" regression as in the "normal population" model and it was even more significant (at the 2 as opposed to the 7 percent level). This finding is consistent with the hypothesis that chronic illness may increase one's predisposition to ozone-related respiratory restricted activity days.

Earlier in this chapter it was hypothesized that information about air quality conditions might influence estimated dose-response relationships. Such information might alert people to poor air quality, with the consequence that they avoid it (and thus make it difficult to observe any adverse effects of air pollution on health); or, it might actually make them aware of minor ailments they might have otherwise overlooked (thus strengthening the observed relationship between air pollution and morbidity). As a test of this hypothesis, we included the dummy variable PSIY1N0 in a regression explaining restricted activity days due to respiratory disease. As noted earlier, it takes on the value unity if air quality as measured by the Pollutant Standard Index was reported in an SMSA in 1979, zero if not.

The results are reported in equation (75), which can be contrasted with equation (49) in Table 4-7. Note that the coefficient on PSIY1N0 was positive (there are more respiratory RADs in SMSAs where the PSI was reported than in those where it was not, all other things being equal) and significant at the 10 percent level. This suggests that people may be more inclined to report feeling poorly if they are aware that air quality is poor. We cannot reject the possibility, however, that the dummy variable

PSIY1N0 was picking up other unobserved differences between SMSAs that influence health but that were not controlled for with the regressors we include. Ozone remained positively and significantly related to respiratory RADs in this equation.

All of the analysis to this point has been based on various samples of individuals who live no more than twenty miles from the nearest ozone, sulfate, particulate and other types of monitors for which data are required in a particular regression. It is interesting to analyze a group of individuals living even closer to the nearest air pollution monitors. (If individuals stay home all the time, one would want monitoring data from as close to individuals' residences as possible.) Accordingly, two equations were estimated using individuals living no more than ten, and then five, miles from the nearest monitor of a particular type.

The results are presented in equations (76) and (77), respectively. Although ozone was significant in this specification of the model using the twenty mile cut-off (see equation (49)), it was not significant in (76) or (77). A number of other variables significant in the earlier regression also lost their significance when the more restrictive distance requirement was imposed. However, because of the F-statistic in equation (77), which imposes the five mile restriction, one cannot maintain with any reasonable degree of confidence that the entire set of regressors is capable of explaining any of the variation in the dependent variable.

Several additional sensitivity analyses were conducted using different measures of ozone during the two-week reference period. They are reported in Table 4-11. Equation (78) again reproduces the basic results presented

Table 4-11. More Sensitivity Analysis of Adult Restricted Activity: Respiratory Disease

Model #	(78)	(79)	(80)	(81)	(82)	(83)	(84)	(85)
Dep. Var. Indep. Var.	TRADRSP	TRADRSP	RADRSPPS (LOGIT)	RADRSPPS (LOGIT)	RADRSPPS (LOGIT)	TRADRSP	TRADRSP	TRADRSP
INTERCEPT	.762 (2.03)	.586 (1.33)	-3.427 (2.07)	-3.201 (2.06)	3.200 (2.08)	-.0487 (0.10)	.0577 (0.10)	-.0657 (0.14)
O3NR01	--	--	--	5.549 (1.13)	8.947 (1.44)	2.530 (2.46)	2.277 (1.88)	--
O3NR01SQ	--	--	--	--	--	--	--	10.143 (1.90)
O3NR03	1.152 (2.60)	--	--	--	--	--	--	--
O3NR04	--	1.547 (1.01)	--	--	--	--	--	--
O3NL01	--	--	-.557 (0.14)	--	--	--	--	--
S4NR01	-.00359 (1.14)	-.00287 (0.91)	-.0119 (0.77)	--	-.0103 (0.73)	-.00135 (0.32)	-.00261 (0.49)	-.000747 (0.19)
SPNR01	-.000153 (0.27)	-.0000179 (0.03)	-.00145 (0.50)	-.00169 (0.61)	--	-.000557 (0.74)	-.000349 (0.33)	-.000512 (0.69)
N2NR01	--	--	--	--	--	--	.000443 (0.75)	--
CONR01	--	--	--	--	--	--	.00433 (0.43)	--
O3S4	--	--	--	-.222 (0.74)	--	--	--	--
O3SP	--	--	--	--	-.0564 (1.14)	--	--	--
RACEW1D0	.105 (2.05)	.104 (2.01)	.448 (1.83)	.470 (1.84)	.442 (1.81)	.107 (1.60)	.122 (1.56)	.110 (1.66)
SEXM1F0	.0335 (1.07)	.0413 (1.13)	-.277 (1.55)	-.112 (0.67)	-.111 (0.56)	.0270 (0.54)	.00834 (0.14)	.0259 (0.52)

Table 4-11 (Continued). More Sensitivity Analyses of Adult Restricted Activity: Respiratory Disease

Model #	(78)	(79)	(80)	(81)	(82)	(83)	(84)	(85)
Dep. Var.								
Indep. Var.	TRADRSP	TRADRSP	RADRSPPS (LOGIT)	RADRSPPS (LOGIT)	RADRSPPS (LOGIT)	TRADRSP	TRADRSP	TRADRSP
MARYIND	-.0159 (0.42)	-.0168 (0.44)	-.284 (1.62)	-.338 (2.02)	-.339 (2.02)	-.0404 (0.81)	-.0115 (0.20)	-.0379 (0.76)
INCOMCON	-.00000476 (2.40)	-.00000466 (2.34)	-.0000138 (1.46)	-.0000134 (1.48)	-.0000134 (1.48)	-.00000451 (1.61)	-.00000667 (2.01)	-.00000447 (1.69)
FAT	-.604 (2.31)	-.603 (2.30)	-.952 (0.91)	-.979 (0.99)	-.982 (0.99)	.167 (0.50)	.194 (0.49)	.171 (0.52)
FATSQ	.120 (2.39)	.120 (2.37)	.159 (0.87)	.182 (0.98)	.182 (0.98)	-.0396 (0.62)	-.0456 (0.59)	-.0402 (0.63)
AGE	.00957 (1.73)	.00990 (1.79)	.0403 (1.43)	.0381 (1.45)	.0386 (1.47)	.0129 (1.72)	.0137 (1.53)	.0128 (1.70)
AGESQ	-.000108 (1.85)	-.000112 (1.91)	-.000548 (1.77)	-.000529 (1.94)	-.000533 (1.86)	-.000156 (1.97)	-.000159 (1.69)	-.000154 (1.95)
SMOKYIND	.0391 (1.12)	.0369 (1.06)	.322 (2.02)	.255 (1.67)	.254 (1.66)	.0242 (0.52)	.0222 (0.40)	.0239 (0.51)
EDCOMCON	.00215 (0.39)	.00199 (0.35)	.0325 (1.14)	.0111 (0.42)	.0106 (0.40)	.00229 (0.30)	-.00263 (0.29)	.00205 (0.27)
CHRLPOUM	.249 (5.33)	.247 (5.29)	.552 (2.79)	.563 (2.96)	.569 (2.99)	.259 (4.36)	.271 (3.79)	.256 (4.31)
AVMAXTMP	-.00452 (3.66)	-.00321 (2.71)	-.0156 (2.84)	-.0169 (3.28)	-.0182 (3.47)	-.00526 (3.22)	-.00602 (3.11)	-.00387 (2.74)
AVPRECIP	.118 (0.69)	.0742 (0.43)	-.0168 (0.00)	.135 (0.17)	.0325 (0.14)	.168 (0.71)	-.0760 (0.26)	.142 (0.60)
HUMIDPF	.00163 (0.77)	.00213 (1.03)	.0274 (2.51)	.0249 (2.33)	.0244 (2.37)	.000400 (0.14)	-.0000595 (0.02)	.000484 (0.17)
R2	.0149	.0134	--	--	--	.0128	.0150	.0121
F or X2	4.34	3.97	65.46	62.53	63.55	2.54	2.03	2.40
N	4,939	4,793	4,637	4,399	4,399	3,355	2,555	3,355

Reproduced from best available copy.

in (49), with the exception that the highest single hourly ozone reading during the two-week reference period was used to characterize exposure to ozone rather than the average of the fourteen daily high readings. The coefficient on ozone in this equation increased in significance suggesting that single ozone "spikes" may play a role in precipitating respiratory illness. In equation (79), the average daily maximum 1-hour reading was replaced with the average hourly concentration over the 336 hours. The coefficient on ozone was no longer significant. This is consistent with the belief that peak exposures to ozone are more likely to induce acute illness than constant exposure to lower levels.

The health data used in this report do not lend themselves well to an analysis of possible lagged effects of ozone and other air pollutants. This is because individuals were asked in the HIS to report only the total number of days of acute morbidity during the two weeks prior to their interview. The actual date on which a restricted activity day, work or school loss day, or bed disability day occurred was not given. This makes it very difficult to select an appropriate lag structure to test for delayed effects.

Nevertheless, in equation (80) an attempt was made to identify a lagged ozone effect.[12] In that logit regression ozone was measured by the average daily 1-hour maximum ozone concentration during the two-week reference period prior to that for which the individual provided data on acute morbidity. Thus, if an individual's reference period was July 15-29 in 1979, ozone was measured in (80) by the average daily maximum at the monitor nearest the individual's home during the two-week period July 1-

July 14. The problems this creates are clear. A very high reading on July 1, for example, would influence the average daily maximum calculated over July 1-14. This average would in turn be used to explain acute morbidity that may have taken place as late as July 29, nearly a full month after the peak episode. In view of this problem, the results are not surprising. The coefficient on ozone was statistically indistinguishable from zero. The signs of the smoking dummy and other independent variables were comparable to those in equation (52) which is identical to (80) save for the ozone measure. This crude attempt does not suggest that lagged effects are unlikely or unimportant. Rather, it demonstrates that a data base built on the HIS is not the best tool with which to investigate such effects.

The evidence for interactive or synergistic effects between ozone and sulfate (Equation (81)) and ozone and suspended particulates (Equation (82)) was not significant. A number of other possible interactive effects were explored but are not reported here. In none of the cases were interactive effects a significant determinant of respiratory-related restricted activity days.

As discussed above, health and other information about individuals not at home during the HIS interview was provided by their spouse or another adult present for the interview. Such information might be more subject to error than if provided directly by the individual. This may be especially true for cause-specific acute morbidity where information about particular disease categories is being sought. This suggests the analysis of respiratory-related restricted activity days might be sensitive to a split in the sample between self-respondents and others.

Several "basic" regressions examining respiratory RADs--numbers (49) - (51)--were reestimated using as observations only those individuals who answered for themselves during the HIS. Equations (83) - (85) correspond to (49) - (51) above, with the exception of a slight change in the way temperature is measured. The self-respondents sample was about 30 percent smaller than the corresponding regression when others were permitted to respond for absent members. In each of the three cases, restricting the analysis to self-respondents increased the coefficient on the ozone variable by about 50 percent and added to its significance as well. The coefficients on ozone in (83) - (85) are significant at the 10 percent level or higher. The association between nitrogen dioxide and respiratory RADs in adults was less significant--compare (84) with (50)--but the earlier finding about ozone and these respiratory RADs was strengthened by the respondents-only analysis.

A final sensitivity analysis concerns potential biases introduced by errors in either air monitoring or the characterization of exposure based on the monitor nearest an individual's home. Both can result in individuals' actual exposures being different, and perhaps significantly so, than those assigned them using the procedure in this study. The errors-in-variables problem, well known in statistics, can lead to biased and inconsistent parameter estimates in ordinary least squares regressions.

To test the sensitivity of our results to possible errors in characterizing individual exposure, the following Monte Carlo procedure was adopted. First, for one of the respiratory RAD models, each individual in the sample had his or her ozone data randomly disturbed by adding or subtracting to it a term drawn from a distribution with mean zero and standard deviation equal to that of the "true" ozone data for the 4,800 individuals; negative values were set equal to zero. Then a regression of respiratory RADs on the randomly disturbed ozone variable and the other explanatory variables was run. This was done fifty times, and the resulting coefficients were averaged. (Results are not presented here but are available from the authors.) The mean estimated ozone coefficient was 0.6221 which is substantially smaller than the "true" coefficient of 1.5926 estimated for the undisturbed model (this is the downward bias one would expect).[13] This underscores the sensitivity of parameter estimates to the measurement of air pollution and other independent variables. It suggests that the introduction of reliable personal monitoring devices may significantly alter estimates of the dose-related effects of ozone and other air pollutants on the population.

The sensitivity analyses above take up more space than we would prefer. But it is important to examine the robustness of the earlier finding of a positive and, generally, statistically significant association between average daily maximum ozone concentrations and respiratory related restricted activity days. These findings did not change appreciably when account was taken of possible seasonal variations in ozone and other pollutants, interactive effects, respondent status, confounding due to pollen, data completeness requirements, informational cues, sensitive populations, or other effects. Lagged effects and distance-to-nearest-monitor variants were also run although in both cases the tests were far from ideal. Overall, there appears to be a positive and often significant association between average daily maximum 1-hour concentrations of ozone and the incidence among adults of acute respiratory illness resulting in restrictions in activity.

4.2.2 Children

Analysis of acute respiratory disease in children is also broken into three categories--restricted activity days, school loss days (for children six years of age or older), and bed disability days. In each case, several different combinations of pollutants are considered using both ordinary least squares and logit techniques.

Equations (86) - (89) in Table 4-12 explore the determinants of respiratory-related restricted activity days in children. Although its sign was always positive, in neither combination of air pollutants or functional forms was ozone significantly associated with respiratory RADs. In (86) the chronic limitation dummy and the temperature variable were the only significant regressors. In (87) nitrogen dioxide was positive and significant at the 10 percent level, a finding consistent with the earlier "all causes" analysis. In addition age and its square were significantly associated with respiratory RADs in (87). Equations (88) and (89) replicate the two previous models using logit regression. About 7 percent of the children in the sample had at least one respiratory RAD during the two weeks. The air pollution variables behaved the same in the alternative functional form, although carbon monoxide did enter negatively and

Table 4-12: Children's Restricted Activity, School Loss, and Bed Disability: Respiratory Disease

Model #	(86)	(87)	(88)	(89)	(90)	(91)	(92)	(93)
Dep. Var.								
Indep. Var.	TRADRSP	TRADRSP	RADRSPPS (LOGIT)	RADRSPPS (LOGIT)	TWSLRSP	TWSLRSP	WSLRSPPS (LOGIT)	WSLRSPPS (LOGIT)
INTERCEPT	.817 (5.19)	.896 (4.71)	-2.109 (3.83)	-1.799 (2.85)	.386 (1.42)	.434 (1.38)	-3.940 (3.11)	-4.213 (2.90)
D3HR01	.441 (0.60)	.213 (0.25)	2.462 (0.99)	1.049 (0.36)	1.951 (1.68)	1.855 (1.29)	9.067 (1.90)	9.117 (1.67)
S4NR01	-.00152 (0.57)	-.00107 (0.32)	-.00256 (0.28)	-.00640 (0.52)	.000693 (0.24)	.00148 (0.34)	.0107 (0.88)	.0219 (1.35)
SPNR01	.003273 (0.53)	-.000703 (0.33)	.00155 (0.84)	.00159 (0.73)	-.000510 (0.89)	-.000877 (1.15)	-.00249 (3.92)	-.00311 (0.97)
N2NR01	--	.000732 (1.80)	--	.00247 (1.83)	--	.000350 (0.73)	--	.00177 (0.95)
CONR01	--	-.00873 (1.29)	--	-.0445 (1.74)	--	-.00773 (0.93)	--	-.0459 (1.22)
RACEWIBO	.0248 (0.62)	.0255 (0.54)	.424 (2.74)	.403 (2.32)	.0197 (0.44)	.0110 (0.23)	.592 (2.50)	.644 (2.39)
SEXMIPD	.00459 (0.15)	.0151 (0.42)	-.244 (2.39)	-.176 (1.63)	-.0362 (1.09)	-.0137 (0.45)	-.389 (2.57)	-.272 (1.58)
INCOMCON	-.00000220 (1.26)	-.00000260 (1.24)	-.00000547 (0.84)	-.00000703 (1.90)	-.00000408 (2.09)	-.00000465 (1.85)	-.00000512 (0.70)	-.00000376 (0.39)
AGE	-.0159 (1.41)	-.0309 (2.14)	-.00750 (0.17)	-.0319 (1.38)	.0135 (0.32)	-0.134 (0.25)	.295 (1.49)	.221 (0.93)
AGESC	.000606 (0.85)	.00147 (1.74)	-.000114 (0.00)	.00270 (0.95)	-.000499 (0.26)	.000739 (0.32)	-.0124 (1.41)	-.00902 (0.89)
CHRLMDUM	.642 (8.35)	.569 (7.11)	1.106 (5.71)	1.065 (4.72)	.300 (5.39)	.674 (6.49)	1.094 (4.42)	1.170 (4.10)
AVMAXTMP	-.00531 (4.79)	-.00621 (4.65)	-.0227 (6.20)	-.0241 (5.69)	-.00279 (2.04)	-.00242 (1.33)	-.0127 (2.12)	-.0126 (1.81)
AVPRECIP	.0483 (0.31)	.0136 (0.46)	.402 (0.75)	.599 (1.13)	-.0107 (0.04)	.0230 (0.15)	.357 (.047)	.698 (0.81)
HUMIDRC	-.00342 (1.77)	-.00325 (1.36)	.00902 (1.30)	.00665 (0.90)	-.00195 (0.91)	-.00147 (0.51)	-.00205 (0.20)	.00487 (0.41)
R2	.0183	.0218	--	--	.0195	.0261	--	--
F or X2	9.00	6.63	101.49	79.62	4.48	3.76	40.78	35.09
N	5,622	4,186	5,622	4,186	2,717	1,977	2,717	1,977

Table 4-12 (continued)

Model #	(94)	(95)	(96)	(97)
Dep. Var. / Indep. Var.	TREDRSP	TREDRSP	BEDRSPFS (LOGIT)	BEDRSPFS (LOGIT)
INTERCEPT	.606 (5.72)	.712 (5.39)	-2.441 (3.69)	-1.356 (2.51)
O3NR01	.932 (2.02)	1.026 (1.71)	4.230 (1.43)	4.414 (1.37)
S4NR01	-.00101 (0.55)	-.00141 (0.39)	-.000325 (0.00)	-.00426 (0.28)
SPNR01	.000255 (0.77)	.0000324 (0.08)	.00149 (0.67)	.000337 (0.14)
N2NR01	--	.000215 (0.77)	--	.00286 (1.79)
CONR01	--	-.00602 (1.28)	--	-.0385 (1.27)
RACEWIRO	.00432 (0.19)	.0203 (0.64)	.451 (2.41)	.511 (2.36)
SEXMIFO	-.00494 (0.24)	.00261 (0.10)	-.232 (2.27)	-.166 (1.13)
INCOMCON	-7.535E-7 (0.64)	-5.244E-7 (0.16)	-.00000431 (0.66)	-.00000519 (0.73)
AGE	-.00675 (0.84)	-.0181 (1.91)	.0215 (0.41)	-.0351 (0.50)
AGESQ	.000236 (0.50)	.000333 (1.42)	-.000904 (0.30)	.00219 (0.64)
CHRLMDUM	.371 (7.15)	.461 (7.08)	1.339 (6.45)	1.337 (5.55)
AVMAXTMP	-.00437 (5.95)	-.00525 (5.67)	-.0229 (5.22)	-.0267 (5.36)
AVPRECIP	.0790 (0.74)	.0397 (.074)	.437 (0.75)	1.194 1.65)
HUMIDRC	-.00332 (3.05)	-.00405 (2.44)	.00213 (0.26)	-.00229 (0.22)
R2	.0151	.0216	--	--
F or X2	7.56	6.55	77.59	69.44
N	5,622	4,185	5,622	4,186

Reproduced from best available copy.

significantly (10 percent level) in (89), for which there is no plausible explanation.

Equations (90) - (93) explore school loss in children, the first two using ordinary least squares and the second two using logit techniques. While ozone was positively related to respiratory school loss when entered in combination with sulfates and total suspended particulates, it was not significant in (91) which includes all five air pollutants. The chronic limitation variable and income were significantly related to children's school loss, the latter suggesting that respiratory RADs are inversely related to annual household income. Ozone was positively and significantly related to respiratory school loss in both logit runs. In those equations, the race dummy became significant and indicates that black children experienced less respiratory school loss days than white children.

Equations (94) - (97) complete the analysis of children's acute respiratory illness. They deal with bed disability days, using both ordinary least squares and logit. In (94) ozone was positively and significantly related to bed disability days when entered in combination with sulfates and particulates. Although its significance was diminished when other pollutants were introduced in (95), the coefficient on ozone remained significant at the 10 percent level. In the case of bed disability, the logit analysis paints a slightly different picture. In (96) and (97) the coefficient on ozone was not significant at the 10 percent level. Thus, ozone may play a role in explaining the number of bed disability days a child experiences, but be a less important determinant of whether a child will have one at all. In the logistic specification of children's respiratory bed disability, nitrogen dioxide entered positively and significantly at the 10 percent level. It was not significant in the OLS specification.

Using restricted activity days, school loss days, or bed disability days, the analysis of acute respiratory disease in children was dominated by the chronic limitation variable and the average maximum temperature. The latter suggests that the higher the average maximum temperature, the fewer the number (or the less likely the occurrence) of days of respiratory disease. The importance of the chronic limitation dummy is clear. The effect of maximum temperature is less so. One reason for its significant negative effect may have to do with colds or respiratory infections. Were these illnesses more common in colder northeastern and midwestern areas, then average maximum temperature would take the sign it has in these regressions. This suggests the desirability of additional work along the lines presented here, but in which more detailed analysis can be conducted of possible epidemic effects.

4.3 Chronic Respiratory Disease

For several reasons, chronic respiratory disease (CRD) and its possible relationship to ozone and other air pollutants is of special interest in this study. First, unlike the analysis of acute morbidity (where clinical studies can provide valuable information), epidemiological studies are the only direct source of information about the possible deleterious long-term effects of exposures to ozone on humans. Also, chronic illness is in several ways more serious than acute. Even some serious acute illnesses appear to be reversible. However, CRD by definition implies a continuing state of ill health which might be moderated but not easily "cured."

Furthermore, chronic illness may predispose one to acute morbidity. In the analysis of acute morbidity above, the dummy variable indicating the presence of chronic illness was always positively and significantly associated with acute illness. Thus, even if ozone were not a direct cause of illness in healthy individuals, it might still have an important indirect effect if it increased the likelihood of chronic illness.

Finally, unlike the case of ozone and acute illness, there are virtually no conclusive studies linking long-term exposures to ozone to chronic health impairments.[13a] This has been ascribed in part to inadequate data and methodology. It is important to see, therefore, whether the larger sample size, the matching of individuals to nearest monitors, the creation of longer term pollution exposure profiles, and the inclusion of residential mobility data--all innovations in this report--alters the assessment of ozone and its relation to chronic morbidity.

In the following analysis the focus is no longer on a particular two-week reference period, but rather on whether or not an individual has CRD that is correlated with long-term exposure to ozone and other air pollutants. Accordingly, air pollution measures are specific not to two-week periods but rather to longer periods of time. In the first set of models, annual average pollution data for 1979 were used; subsequently, analyses were conducted using multi-year average concentrations that may begin to characterize individuals' long-term exposure to ozone and other air pollutants.

4.3.1 Adults

With the exception of the period over which air pollution data are summarized, the basic model used to explore the determinants CRD is similar to that used above. It includes socioeconomic and personal characteristics, meteorological information, and certain other variables as well. Ozone, sulfur dioxide, carbon monoxide, and nitrogen dioxide were summarized by their 1979 hourly concentrations. Sulfates and total suspended particulate matter were summarized using 1979 twenty-four hour readings. The dummy variable indicating the presence of a chronic illness was excluded from the list of regressors, since the respiratory component thereof is what we are trying to explain.

Equations (1) and (2) in Table 4-13 present the first analyses of the determinants of CRD. Ozone was measured by the average daily 1-hour maximum reading over the year 1979, as were carbon monoxide and nitrogen dioxide. Sulfates and total suspended particulates were measured by average daily (24-hour) concentrations. The sample used in (1) included 7,210 adults, 537 (7.5 percent) of whom reported having CRD. The sample used in (2)--which is smaller because more air pollution data were required--included 5,891 individuals, of whom 435 (7.4 percent) had CRD. In neither equation did any of the air pollutants exert a positive and significant effect on the likelihood of CRD. Income was negative and highly significant, while education was positively and significantly

Table 4-18a. Chronic Respiratory Disease Among Adults: 1979 Pollution Data

Model #	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. Var. / Indep. Var.	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)
INTERCEPT	-3.399 (3.05)	-3.056 (2.41)	-3.240 (2.94)	-2.854 (2.28)	-3.265 (2.99)	-2.870 (2.31)	-3.329 (3.00)	-2.949 (2.34)
O3NA01	.332 (0.37)	.411 (0.14)	.897 (0.36)	.343 (0.10)	-- --	-- --	-- --	-- --
O3NA03	-- --	-- --	-- --	-- --	-- --	-- --	-.122 (0.17)	-.357 (0.37)
O3NA04	-- --	-- --	-- --	-- --	2.136 (0.47)	.125 (0.00)	-- --	-- --
SPNA01	.00134 (0.58)	.00151 (0.62)	.00134 (0.58)	.00166 (0.65)	.00137 (0.60)	.00168 (0.66)	.00148 (0.65)	.00165 (0.64)
S4NA01	.00564 (0.35)	.0150 (0.69)	.00719 (0.37)	.0156 (0.72)	.00772 (0.41)	.0159 (0.74)	.00952 (0.48)	.0177 (0.81)
CONA01	-- --	-.0416 (1.60)	-- --	-.0419 (1.62)	-- --	-.0421 (1.61)	-- --	-.0431 (1.66)
N2NA01	-- --	.00113 (0.69)	-- --	.00116 (0.71)	-- --	.00122 (0.78)	-- --	.00156 (0.90)
RACEWIBO	.145 (0.98)	.124 (0.76)	.133 (0.93)	.114 (0.70)	.136 (0.92)	.115 (0.71)	.142 (0.96)	.116 (0.72)
SEXMIFO	-.149 (1.46)	-.103 (0.91)	-.164 (1.52)	-.122 (1.09)	-.144 (1.62)	-.122 (1.09)	-.163 (1.61)	-.122 (1.09)
MARYINO	-.0970 (0.95)	-.0539 (0.45)	-.0545 (0.56)	-.0146 (0.14)	-.0546 (0.56)	-.0143 (0.14)	-.0541 (0.55)	-.0132 (0.10)
INCUMCON	-.0000278 (4.36)	-.0000321 (5.14)	-.0000270 (4.98)	-.0000310 (5.03)	-.000270 (4.94)	-.0000310 (5.03)	-.0000270 (4.87)	-.0000309 (5.02)
FAT	-.598 (1.14)	-.815 (1.22)	-.511 (1.00)	-.714 (1.09)	-.510 (1.09)	-.714 (1.08)	-.505 (0.99)	-.712 (1.07)
FATSQ	.169 (1.47)	.196 (1.54)	.155 (1.37)	.192 (1.49)	.157 (1.37)	.193 (1.49)	.156 (1.36)	.183 (1.47)

Table 4-13 (continued) Chronic Respiratory Disease Among Adults: 1979 Pollution Data

Model #	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. Var. Indep. Var.	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)
AGE	.0163 (1.13)	.0208 (1.24)	-- --	-- --	-- --	-- --	-- --	-- --
AGESQ	-.0000654 (0.42)	-.000123 (0.71)	.000100 (3.40)	.0000900 (3.04)	.000100 (3.81)	.0000900 (3.04)	.000100 (3.81)	.0000901 (3.04)
NCIGSDYN	.00174 (0.46)	.00392 (0.96)	.00250 (0.58)	.00489 (1.23)	.00249 (0.63)	.00483 (1.23)	.00248 (0.67)	.00490 (1.23)
EDCUMCON	.0453 (2.91)	.0526 (3.06)	.0470 (3.04)	.0546 (3.18)	.0471 (3.04)	.0546 (3.18)	.0470 (3.03)	.0544 (3.17)
TEMPAN	.00439 (0.57)	.00227 (0.22)	.00448 (0.57)	.00245 (0.24)	.00477 (0.65)	.00271 (0.28)	.00596 (0.76)	.00390 (0.39)
PRECAN	-.00319 (0.43)	-.00436 (0.88)	-.00319 (0.43)	-.00428 (0.86)	-.00332 (0.83)	-.00435 (0.88)	-.00391 (0.97)	-.00465 (0.93)
HUMIDAN	.00942 (1.16)	.00542 (0.61)	.00926 (1.14)	.00528 (0.53)	.00930 (1.15)	.00538 (0.61)	.0101 (1.23)	.00591 (0.66)
F or X2 N	66.08 7,210	62.32 5,891	64.86 7,210	60.76 5,891	64.95 7,210	60.75 5,891	64.76 7,210	60.89 5,891

related to the presence of CRD. This latter effect was unexpected since more educated individuals, by being better informed about health matters, might be expected to have less illness. It may be that these individuals are more likely to have had CRD diagnosed as such, than those with less education and, perhaps, less access to medical attention.

One puzzling aspect of equations (1) and (2) is the performance of the age measures. One would expect older individuals to have more CRD. One explanation for the insignificant coefficients might be the presence of collinearity between age and its square. Diagnostic tests for collinearity revealed this to be a problem. The regressions were rerun using the square of age as a single measure of its effect (because a non-linear relationship was suspected). These results are presented in (3) and (4). The square of age joined education and income as significant correlates of CRD.

Two additional measures of ozone exposure were tested. In (5) and (6) ozone was measured by its annual average hourly concentration in 1979. The results were similar to (3) and (4). In equations (7) and (8) the regressions were rerun using the single highest hourly reading for the year as the measure of exposure to ozone. As in the analysis based on annual average daily maximum ozone readings, income, education, and age were the only significant determinants of CRD. The fact that the number of cigarettes smoked per day was not significant is unsettling. Later analysis sheds some light on this matter.

It should not be surprising that no significant relationship appeared to exist between CRD and ozone and other air pollutants as characterized by 1979 values. One year's exposure to even high pollution levels might not be sufficient to cause or contribute to the onset of CRD. Annual average values for a single year, 1979 in equations (1)-(8), are used in some studies with the implicit understanding that they are proxies for multi-year exposures. If the values for a particular year are atypical, spurious correlations can arise.

To minimize the likelihood of this occurrence, an air pollution data set was created consisting of annual average pollution data for as many as six years, 1974-1979. These data were assembled on a monitor-by-monitor basis and matched to individuals, thus enabling us to assign to individuals longer-term average air pollution concentrations than have been used in previous studies. This sheds light on the bias that may result from using 1979 data alone to explain CRD.

Equations (9) and (10) (in Table 4-14) present the first analysis of CRD in adults using the multi-year averaged air pollution data. The sample in (9) consisted of 7,240 adults including 555 reporting CRD. Six thousand and five adults form the sample in (10), 468 of whom suffer from CRD. In the two equations all pollutants but one were measured by their (multi-year) arithmetic mean. Total suspended particulates were measured by (multi-year) geometric mean concentrations. Because no data were readily available for sulfates for the years 1974-1978, sulfur dioxide data were used instead.

The introduction of air pollution data averaged over longer periods of time made a considerable difference in the results. In equation (9) ozone was significantly associated with CRD at the 10 percent level. Sulfur dioxide also entered positively and was significant at the 5 percent level,

Table 4-14. Chronic Respiratory Disease Among Adults: Multi-year Pollution Data and Sensitivity Analysis

Model #	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Dep. Var.	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)
Indep. Var.								
INTERCEPT	-4.125 (3.46)	-4.229 (2.96)	-4.192 (3.51)	-4.152 (2.93)	-4.735 (3.05)	-5.210 (2.76)	-4.642 (2.81)	-2.92 (2.17)
O3ANARAV	13.341 (1.53)	7.555 (0.75)	-- --	-- --	7.677 (0.66)	-3.761 (0.28)	13.556 (1.23)	22.485 (2.30)
O3ANMAX2	-- --	-- --	-1.307 (1.23)	-1.223 (0.77)	-- --	-- --	-- --	-- --
SPANGEAV	.00101 (0.28)	.00285 (0.71)	.00351 (0.38)	.00397 (1.05)	.00479 (0.99)	.00749 (1.47)	-.000115 (0.00)	.00275 (0.66)
S2ANARAV	.00657 (2.04)	.00994 (2.52)	.00703 (2.12)	.00972 (2.50)	.00829 (1.93)	.0102 (1.98)	.0100 (1.82)	.00330 (0.90)
COANARAV	-- --	-.0757 (1.14)	-- --	-.0784 (1.18)	-- --	-.0269 (0.80)	-- --	-- --
N2ANARAV	-- --	-.00243 (0.65)	-- --	-000223 (0.00)	-- --	-.00398 (0.30)	-- --	-- --
O3S2	-- --	-- --	-- --	-- --	-- --	-- --	-- --	-- --
RACEWIBO	.200 (1.38)	.155 (0.99)	.202 (1.39)	.158 (1.01)	.492 (2.27)	.478 (2.03)	.348 (1.98)	.0718 (0.45)
SEXMIFO	-.155 (1.56)	-.197 (1.82)	-.155 (1.56)	-.197 (1.82)	-.0350 (0.28)	-.0548 (0.39)	-.160 (1.34)	-.0633 (0.55)
MARYIND	-.0577 (0.60)	-.0124 (0.10)	-.0559 (0.57)	-.00833 (0.10)	-.0456 (0.35)	.0123 (0.10)	-.0452 (0.39)	-.0417 (0.37)
INCOMCON	-.0000245 (4.49)	-.0000259 (4.34)	-.0000244 (4.46)	-.0000259 (4.33)	-.0000315 (3.99)	-.0000282 (3.34)	-.0000227 (3.50)	-.0000208 (3.16)
FAT	-.363 (0.52)	-.101 (0.14)	-.391 (0.63)	-.107 (0.14)	-.235 (0.32)	.251 (0.30)	-.0203 (0.00)	-.249 (0.35)

Reproduced from
best available copy.

Table 4-14 (continued) Chronic Respiratory Disease Among Adults: Multi-year Pollution Data and Sensitivity Analysis

Model #	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Dep. Var. Indep. Var.	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)
FATSQ	.111	.0676	.113	.0696	.106	.0173	.0406	.0693
	(0.94)	(0.52)	(0.96)	(0.53)	(0.77)	(0.10)	(0.26)	(0.50)
AGE	--	--	--	--	--	--	.00615	--
	--	--	--	--	--	--	(2.00)	--
AGESQ	.0000789	.0000777	.0000782	.0000773	.000595	.0000413	--	.0000659
	(3.01)	(2.72)	(2.98)	(2.70)	(1.57)	(1.01)	--	(2.19)
NCIGSDYN	.00187	.000976	.00180	.000927	-.00586	-.00480	-.0000847	.00382
	(0.51)	(0.24)	(0.49)	(0.22)	(1.06)	(1.45)	(0.00)	(0.94)
SMOKYINO	--	--	--	--	--	--	--	--
	--	--	--	--	--	--	--	--
EDCOMCON	.0294	.0276	.0283	.0254	.0366	.0276	.0268	.0101
	(1.93)	(1.66)	(1.96)	(1.59)	(1.77)	(1.23)	(1.50)	(0.57)
TEMPAN	.00416	.00277	.0109	.00446	.0103	.00890	.00796	-.0133
	(0.47)	(0.22)	(1.11)	(0.36)	(0.85)	(0.54)	(0.62)	(1.30)
PRECAN	-.00249	-.00789	-.00484	-.00897	-.00420	-.00373	-.00242	-.00239
	(0.58)	(1.72)	(1.29)	(1.95)	(0.84)	(1.60)	(0.46)	(0.57)
HUMDAN	.0161	.0215	.0184	.0222	.00983	.0195	.0147	.0129
	(1.74)	(2.03)	(2.00)	(2.11)	(0.83)	(1.43)	(0.95)	(1.24)
PGASPIPD	--	--	--	--	--	--	-.533	--
	--	--	--	--	--	--	(1.44)	--
POLLENAN	--	--	--	--	--	--	--	--
	--	--	--	--	--	--	--	--
PCDDCN	--	--	--	--	--	--	--	--
	--	--	--	--	--	--	--	--
F or X2	52.25	44.17	61.14	64.34	55.23	45.92	33.24	39.43
N	7,240	6,005	7,240	6,005	4,109	3,419	4,939	4,386

Table 4-14 (continued) Chronic Respiratory Disease Among Adults: Multi-year Pollution Data and Sensitivity Analysis

Model #	(17)	(18)	(19)	(20)	(21)
Dep. Var.	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)
Indep. Var.					
INTERCEPT	-3.251 (1.99)	-4.692 (3.65)	-4.071 (3.43)	-1.254 (0.69)	-1.556 (0.85)
O3ANARAV	11.932 (1.03)	34.483 (1.77)	14.109 (1.62)	-.846 (0.00)	4.005 (0.26)
O3ANMAX2	-- --	-- --	-- --	-- --	-- --
SPANGEAV	.00543 (1.20)	.000946 (0.25)	.000196 (0.24)	-.0191 (2.11)	-.0143 (1.69)
S2ANARAV	.00735 (1.63)	.0199 (1.69)	.00739 (2.16)	.0143 (2.39)	.00932 (1.77)
COANARAV	-.114 (1.49)	-- --	-- --	-- --	-- --
N2ANARAV	-.00173 (0.40)	-- --	-- --	-- --	-- --
O3S2	-- --	-.641 (1.19)	-- --	-- --	-- --
RACEW130	.0121 (0.00)	.207 (1.42)	.221 (1.53)	.210 (1.00)	.173 (0.82)
SEXMIFO	-.117 (0.93)	-.155 (1.56)	-.147 (1.50)	-.162 (1.13)	-.198 (1.23)
MARYIND	.0159 (0.14)	-.0553 (0.59)	-.0782 (0.31)	.231 (1.83)	.316 (2.06)
INCOMCON	-.0000229 (3.20)	-.0000245 (4.49)	-.0000249 (4.46)	-.0000240 (2.84)	-.0000242 (2.87)
FAT	.109 (0.14)	-.375 (0.60)	-.522 (0.35)	-1.944 (2.03)	-1.953 (2.03)

Table 4-14 (continued) Chronic Respiratory Disease Among Adults: Multi-year Pollution Data and Sensitivity Analysis

Model #	(17)	(18)	(19)	(20)	(21)
Dep. Var. / Indep. Var.	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)
FATSQ	.00765 (0.00)	.110 (0.93)	.134 (1.15)	.418 (2.33)	.418 (2.33)
AGE	-- --	-- --	.00321 (3.24)	-- --	-- --
AGESQ	.0000744 (2.28)	.0000796 (3.03)	-- --	.0000375 (0.91)	.0000306 (0.75)
NCIGSDYN	.00396 (0.88)	.00187 (0.51)	-- --	-.00384 (0.64)	-.00338 (0.57)
SMOKYIND	-- --	-- --	.0307 (0.33)	-- --	-- --
EDCOMCON	.0104 (0.54)	.0298 (1.96)	.0297 (1.38)	.0202 (0.89)	.0205 (0.91)
TEMPAN	-.0127 (0.92)	.00287 (0.32)	.00406 (0.45)	.0391 (2.12)	.0290 (1.56)
PRECAN	-.00304 (1.52)	-.00326 (0.87)	-.00217 (0.50)	-.0216 (2.43)	-.0136 (1.65)
HUMIDAN	.0197 (1.54)	.0193 (2.00)	.0168 (1.83)	.000973 (0.00)	.00955 (0.62)
PGASPIPD	-- --	-- --	-- --	-- --	-- --
POLLENAN	-- --	-- --	-- --	.0707 (2.88)	-- --
PCDDCN	-- --	-- --	-63.585 (0.73)	-- --	-- --
F OR X2 N	47.15 4,146	63.65 7,240	65.57 7,404	34.89 2,969	27.36 2,969

while particulates were not significant. Income and age (squared) were still significantly correlated with CRD, as was education although at a lower level of significance than in (1)-(8). In addition, average annual humidity became significant at the 10 percent level.

The coefficient on ozone in equation (9) was sensitive to the inclusion of additional air pollution variables, as (10) illustrates. When carbon monoxide and nitrogen dioxide were added to the estimating equation, the ozone coefficient was reduced in size and significance. The coefficient on sulfur dioxide increased as a result of this addition, and its significance level also increased. The sex dummy also became significant at the 10 percent level in (10), and indicates that men are less likely than women to have CRD. In equations (11) and (12), exposure to ozone was characterized by the average of the second-highest annual readings for the years for which data are available at the monitor nearest each individual's home. The other pollutants were still measured by their multi-year annual averages. The coefficient on ozone turned negative when this change was made although it was not significant. Since chronic morbidity is less likely than acute illness to be related to extreme values, the performance of O3ANMAX2 is not particularly surprising.

We believe the use of air pollution data averaged over a number of years to characterize long-term exposure is a significant improvement on the use of one year's data. However, analyzing the effect of air pollution on CRD would still be a difficult task even if one had accurate historical air pollution data from all locations. This is because individuals generally do not live in one area their whole lives. Lacking information on residential histories, one might falsely ascribe a high incidence of CRD in a particular area (say St. Petersburg, Florida) to air quality there even though many individuals now living there may have lived most of their lives in industrial cities in the northeast or midwest. This would be a textbook case of spurious correlation.

We made use of the residential mobility supplement to the 1979 HIS in an attempt to circumvent this problem. Using the supplement, we identified a group of adults who had been living at their present address for at least five years at the time of the 1979 HIS, the approximate period for which air pollution data were available. A number of individuals in this group had been at their present address their entire lives. In all we identified 4,108 such individuals of whom 309 (or 7.5 percent) reported having a CRD. The results of some of our analysis using the subsample of these "residentially stable" individuals is presented in equations (13) and (14).

Comparing (9) with (13), it is clear that confining the analysis to relatively long-time residents of the SMSAs in our sample has the effect of reducing both the size and the significance of the ozone variable. Whereas in (9) ozone was significant at the 10 percent level, that was no longer true in (13). Sulfur dioxide remained significantly associated with the likelihood of CRD, although no longer at the 5 percent level. Education, income, age and race were all significant at the 10 percent level or higher. Comparing (10) with (14) results in a similar conclusion.

It is important to explore a number of other potentially confounding influences on the estimated coefficients of ozone. Indoor air pollution is one example. Recall that data were collected on the use of natural gas as a cooking fuel in the thirty-eight largest SMSAs in the country in 1979. After converting the data to percentages of occupied housing units using natural gas, we identified 4,939 individuals for whom we had multi-year air pollution data as well as data on the probability that the household cooked with gas (all households in the same SMSA were assigned the same probability). Equation (15) introduces the variable PGASPIPD which is the probability discussed above.

If equation (15) is compared to (9)--neither are restricted to the residentially stable subset--one can see that the introduction of the gas stove variable did not reduce the size of the coefficient on ozone but did reduce its significance. The gas stove variable also reduced somewhat the significance of the coefficient on sulfur dioxide although it remained significant at the 10 percent level. Equation (15) suggests that the higher the probability of a household's having a gas stove, the less likely is an individual from that household to have CRD. Since natural gas is a source of both nitrogen dioxide and carbon monoxide, this was not an expected result.

These results should be interpreted with caution, however. The gas stove data are aggregated to the SMSA level in contrast to almost every other variable in our data set. Individual data or data disaggregated to the census tract level would be preferred. As concerns the effects of gas stoves on the ozone coefficient, two factors should be considered. First, the sample size varies substantially between (9) and (15)--7,210 individuals vs. 4,939 respectively--because gas stove data are available only for a subset of SMSAs. This difference by itself could account for the change in the ozone coefficient. Also, the SMSAs for which gas stove data are available are the thirty-eight largest in the U.S. Thus, the samples may be quite different with respect to population characteristics and perhaps even other dimensions which affect health but which are not included among the set of regressors. In summary, the effort in (15) to address indoor air pollution should be viewed as preliminary.

Separate regressions were also run on a set of self-respondents although the imprecise recollections of others should be a less serious problem for chronic than for acute respiratory disease--it is unlikely that CRD would escape the attention of a spouse or other related adult answering for an absent household member. These are reported in equations (16) and (17) and can be compared with equations (9) and (10) since neither pair of equations is restricted to the set of residentially stable respondents. When the sample was confined to self-respondents, the coefficients on ozone increased substantially in magnitude and significance. In (16), where ozone was entered with both total suspended particulate matter as well as sulfur dioxide, ozone was positively related to the likelihood of CRD and was significant at greater than the 5 percent level. Note that the coefficient on sulfur dioxide was no longer significant in either regression restricted to self-respondents.

A number of sensitivity analyses were conducted to explore synergistic effects between air pollutants. An example is presented in equation (18), estimated over 7,240 adults of whom 555 reported a CRD. Ozone and sulfur dioxide were both entered individually along with total suspended particulates and an interactive term between ozone and sulfur dioxide.

Both ozone and sulfur dioxide were significant at the 10 percent level by themselves. The interactive effect was negative though insignificant. In other regressions run but not reported the interactive term was never significant when the air pollutants in question were also entered individually. One problem with these regressions was the multicollinearity which arose when a pollution variable was entered by itself and in interaction with another.

It is possible that access to medical care may reduce the likelihood that an individual will contract a CRD. This would be the case, for example, if better access implied earlier detection of developing problems followed by "preventive maintenance." To explore this possibility, the basic model was augmented with a variable measuring the per capita concentration of doctors in the SMSAs from which the observations were drawn (PCDOCN). Equation (19) presents the results for a sample of 7,404 adults, 7.4 percent of whom (or 567) reported having CRD. The coefficient on doctors per capita was negative but insignificant suggesting that the availability of medical care, as measured here, has little effect on the prevalence of CRD. In future research we intend to improve our measures of access to medical care, as well as employ more detailed specifications of the relationship between the supply and demand for physicians and the consequent effects on chronic morbidity.[14]

We also explored the sensitivity of our results to the inclusion of data on ragweed (pollen) concentrations. To do so, the basic model from equation (9) was run with annual average pollen data added. This reduced the size of the sample considerably because pollen data were only available for a subset of SMSAs. The results are presented in (20). Ozone was not significantly associated with CRD in (20) even though it was significant at the 10 percent level in (9). The introduction of pollen had dramatic effects on many of the other signs as well. Variables that were significant in equation (9)--age and education--were no longer significant in (20). Variables that were insignificant in (9)--suspended particulates, the marital status dummy, FAT and its square, and annual temperature--became significant in (20) when pollen was included.

A possible explanation for this change may be the special characteristics of the SMSAs for which pollen data were available--they are the very largest metropolitan areas in the country. As such they may have important characteristics that differentiate them from smaller SMSAs which are not controlled with the other regressors that are included. To determine whether effects attributed to ozone were due instead to pollen, we replicated (9) on the same set of individuals used in equation (20). In other words, CRD was explained both with and without pollen data among the same individuals. The results are presented in (21). Even in the absence of pollen data, ozone was an insignificant determinant of the likelihood of CRD in this subsample. Thus, effects attributed to ozone apparently were not due instead to pollen concentrations.

This exercise does point out an important aspect of the analysis. The estimated coefficient on ozone, while robust in certain sensitivity analyses, can change substantially across other subsets of the data. Analysis to illustrate the presence of influential observations shows both that such observations do exist and, when they exist, they tend to be those observations where CRD is reported. Of course, this is what one would expect. This warns one, however, that when the sample is subset, previously influential observations can be dropped, thus tending to change the estimated parameters. This sensitivity must be taken into account in drawing conclusions from these results.

In reviewing the analyses to this point, one very striking result is that the smoking variable (number of cigarettes per day) was not positively and significantly associated with CRD in any of the nineteen models. All of these regressions were rerun using the dummy variable (SMOKY1NO) that distinguished between those who classified themselves as regular or occasional smokers and those who were former smokers or had never smoked at all. This substitution made no difference: current smoking status was still unrelated to the presence of CRD.

One possible explanation for this puzzling result is that many of those suffering from CRD may have stopped smoking even though they once smoked. This suggests the utility of subdividing non-smokers into two groups: those who used to smoke and those who never smoked. This was done in equation (22) of Table 4-15 where two dummy variables were created identifying current smokers and former smokers (never-smokers are the excluded category). Both variables had positive signs indicating that present and former smokers are more likely than never-smokers to report a CRD. Moreover, the coefficient on the former-smoker dummy variable was highly significant (at nearly the 1 percent level). Thus, it would appear that both NCIGSDYN or SMOKY1NO may be too crude to capture the effects of smoking on CRD. The results indicate that smoking is strongly related to

Table 4-15. Smoking, Air Pollution, and Chronic Respiratory Disease

Model #	(22)	(23)	(24)	(25)	(26)	(27)	(28)	(29)
Dep. Var. Indep. Var.	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)
INTERCEPT	-4.161 (3.51)	-7.446 (3.52)	-2.735 (1.18)	-1.336 (0.67)	-6.570 (2.71)	-2.973 (1.05)	-2.164 (0.88)	-2.804 (5.46)
O3ANARAV	12.465 (1.48)	42.920 (3.22)	-13.743 (1.04)	-1.447 (0.10)	40.021 (2.53)	-23.109 (1.38)	-7.090 (0.40)	15.171 (1.92)
SPANGEAV	.000943 (0.26)	-.00307 (0.51)	-.00583 (0.38)	.03765 (1.30)	-.00474 (0.69)	-.00538 (0.56)	.0119 (1.89)	.00151 (0.55)
S2ANARAV	.00639 (1.98)	.00689 (1.73)	.0130 (2.61)	-.00260 (0.48)	.0133 (2.03)	.0200 (2.33)	-.00162 (0.24)	.00255 (0.93)
CUANARAV	-- --	-- --	-- --	-- --	-.0754 (0.70)	-.119 (0.89)	-.0224 (0.20)	-- --
N2ANARAV	-- --	-- --	-- --	-- --	-.00192 (0.32)	.000264 (0.00)	-.00251 (0.37)	-- --
RACEWHBD	.245 (1.67)	.0120 (0.00)	.365 (0.33)	.446 (1.73)	-.106 (0.47)	.381 (0.90)	.466 (1.73)	-- --
SEXMIFO	-.174 (1.71)	-.419 (2.43)	.260 (1.19)	-.284 (1.65)	-.424 (2.29)	.204 (0.89)	-.392 (2.06)	-.139 (1.62)
SCURRENT	.109 (1.05)	-- --	-- --	-- --	-- --	-- --	-- --	-- --
SFORMER	.202 (2.37)	-- --	-- --	-- --	-- --	-- --	-- --	-- --
MARYIND	-.102 (1.04)	-.0757 (0.51)	.280 (1.22)	-.317 (1.50)	-.0904 (0.49)	.334 (1.36)	-.234 (1.27)	-- --
INCOMCON	-.0000256 (4.67)	-.0000145 (1.77)	-.0000273 (2.26)	-.0000358 (3.70)	-.0000194 (2.11)	-.0000309 (2.40)	-.0000305 (2.93)	-.0000243 (5.15)
FAT	-.425 (0.68)	1.064 (0.87)	-1.159 (0.94)	-1.408 (1.33)	.945 (0.75)	-.703 (0.59)	-1.085 (0.89)	
FATSQ	.114 (0.35)	-.178 (0.74)	.195 (0.57)	.351 (1.75)	-.115 (0.48)	.132 (0.54)	.279 (1.18)	

Table 4-15 (continued) Smoking, Air Pollution, and Chronic Respiratory Disease

Model #	(22)	(23)	(24)	(25)	(26)	(27)	(28)	(29)
Dep. Var. Indep. Var.	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)
SMOKYINØ	-- --	-- --	-- --	-- --	-- --	-- --	-- --	.0261 (0.30)
AGE	.00745 (2.38)	-- --	-- --	-- --	-- --	-- --	-- --	.00844 (3.58)
AGESQ	-- --	.0000476 (1.21)	.0000529 (0.34)	.000126 (2.41)	.0000648 (1.50)	.0000343 (0.57)	.0000982 (1.71)	-- --
EDCOMCON	.0287 (1.88)	.0507 (2.23)	.00725 (0.22)	.00556 (0.20)	.0504 (2.03)	-.00413 (0.10)	.0174 (0.54)	.0231 (1.64)
TEMPAN	.00496 (0.56)	.00324 (0.22)	.0531 (2.12)	-.0131 (1.23)	.00729 (0.35)	.0552 (2.16)	-.0304 (1.46)	-.00728 (0.90)
PRECAN	-.00241 (0.56)	-.0000935 (0.00)	-.0101 (1.30)	-.00136 (0.22)	-.00808 (1.14)	-.0125 (1.32)	-.00350 (0.44)	-- --
HUMIDAN	.0142 (1.55)	.0310 (1.97)	-.0133 (0.81)	.0159 (1.09)	.0244 (1.40)	-.00779 (0.40)	.0342 (1.84)	-- --
F or X2	59.16 7,244	34.09 3,134	25.54 1,432	51.25 2,558	35.78 2,660	24.59 1,244	42.27 2,104	57.19 8,214

the presence of respiratory disease. In future research we intend to go beyond these limited measures to make use of the tar and nicotine data that are part of the smoking supplement in the 1979 HIS.

Differentiating between never, former, and current smokers raises another interesting possibility: might the effects of ozone and other air pollutants on CRD differ among individuals in these different categories? To determine this, we subdivided a sample of 7,244 adults into the three categories: 3,194 never smokers, 1,492 former smokers, and 2,558 current smokers. We then ran our basic regression on each of the three subsamples. The results, presented in equations (23)-(25), are quite interesting. In equation (23), based on a sample of individuals who have never smoked, ozone had a very large positive and highly significant coefficient. This is not the case in either (24) or (25), which used the samples of former smokers and current smokers respectively. In neither of those samples was the multi-year average ozone concentration significantly associated with CRD. Sulfur dioxide was significantly associated with CRD among former smokers as it was among never smokers, but it was not significant in the current smokers regression.

The results of (23)-(25) suggest an intriguing possibility: if individuals elect to begin and continue smoking, they may so increase their likelihood of developing CRD that the marginal effect of exposure to ambient ozone pollution may be very small or non-existent. Among individuals who do not smoke, however, ozone is positively and significantly associated with the likelihood of CRD in the models estimated above.

To explore the sensitivity of these findings, we replicated the experiment in (23)-(25) in a model that includes five pollutants. In this model, first presented in equation (10), ozone was positively but insignificantly related to adult CRD. However, when separate models were estimated for never, former and current smokers, ozone was positively and significantly associated with CRD among the never-smokers (Equation (26)), as it was in the three-pollutant specification. Among former and present smokers, however, the association was not a significant one. Sulfur dioxide was positively and significantly associated with CRD among never and former smokers, while total suspended particulates were positively and significantly related to CRD among current smokers. In all, (26) - (28) strengthen the case that long-term exposure to ozone may be most likely to result in CRD among non-smokers.

Throughout the analysis of CRD, the same basic set of regressors was used. Air pollution was summarized in different ways and several Different measures of smoking status were explored, but little was done with the other independent variables even though a number of them were seldom if ever significantly associated with the likelihood of CRD. Dropping them because they did not perform as expected could have biased the coefficients of the ozone and other air pollution variables of interest.

As a final experiment in the analysis of adult CRD, however, a "lean" version of the basic model was tested. This version excluded a number of the variables that were never significantly associated with CRD in the earlier analysis. The results from this version of the model are presented in equation (29) in Table 4-15. They do not differ substantially from earlier results. The multi-year annual average hourly ozone concentration was positively associated with the likelihood of CRD, at between the 5 and 10 percent level. Neither suspended particulates nor sulfur dioxide were significant. Income was negatively and quite significantly associated with CRD, while age was positive and significant in the same equation.

4.3.2 Children

The analysis of ozone and chronic respiratory disease among children is much more brief than that for adults. This is because the analysis revealed no relationship between ozone and CRD in children and the approach in this study has been to confine sensitivity analyses primarily to those areas where a positive and statistically significant ozone effect was identified. Even had such effects been identified, the sensitivity analysis would have been more limited than that above because less data are available for children. For instance, information about almost all children was provided by a parent or other related adult--thus, no "self-respondents" analysis could be conducted. Also, no smoking data were available for those below the age of 17 nor could accurate household smoking profiles be constructed. This limits the variations on the basic model that could be considered.

The results of the analysis of CRD in children are presented in Table 4-16. Equation (30) is based on a sample of 8,483 children aged 16 and below. Of this group 516 had CRD. The corresponding figures for equation (31) are 6,886 and 429 respectively. Ozone and other air pollutants were measured by 1979 average concentrations in the first two models. Although

Table 4-16. Chronic Respiratory Disease Among Children

Model #	(30)	(31)	(32)	(33)	(34)	(35)
Dep. Var.						
Indep. Var.	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)
INTERCEPT	-3.406 (4.79)	-3.450 (4.07)	-3.463 (4.18)	-4.498 (4.19)	-2.915 (2.24)	-3.973 (2.47)
O3NA01	3.254 (1.21)	.928 (0.30)	-- --	-- --	-- --	-- --
SPNA01	-.00286 (1.22)	-.00608 (2.29)	-- --	-- --	-- --	-- --
S4NA01	.0168 (0.90)	.0135 (0.95)	-- --	-- --	-- --	-- --
CONA01	-- --	.00828 (0.36)	-- --	-- --	-- --	-- --
N2NA01	-- --	.00262 (1.63)	-- --	-- --	-- --	-- --
O3ANARAV	-- --	-- --	.0176 (0.00)	-9.822 (0.92)	-15.349 (1.26)	-18.623 (1.32)
SPANGEAV	-- --	-- --	-.00487 (1.28)	-.00170 (0.39)	.00320 (0.63)	-.000384 (0.00)
S2ANARAV	-- --	-- --	.00689 (2.14)	.00415 (1.08)	.00251 (0.45)	-.000339 (0.00)
COANARAV	-- --	-- --	-- --	.158 (2.45)	-- --	-- --
N2ANARAV	-- --	-- --	-- --	-.00570 (1.56)	-- --	-.00495 (1.02)
PGASPIPD	-- --	-- --	-- --	-- --	.674 (1.74)	.0567 (1.20)
RACEW1B0	.396 (2.89)	.401 (2.67)	.367 (2.74)	.310 (2.16)	.366 (2.36)	.350 (2.17)

Table 4-16 (continued) Chronic Respiratory Disease Among Children

Model #	(30)	(31)	(32)	(33)	(34)	(35)
Dep. Var. Indep. Var.	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)	PRSCHRSP LOGIT)	PRSCHRSP (LOGIT)
SEXMIFO	.360	.385	.353	.363	.283	.321
	(3.89)	(3.79)	(3.82)	(3.57)	(2.55)	(2.71)
INCOMCON	-.00000065	.00000132	.00000405	.00000259	.00000845	.00000159
	(0.14)	(0.22)	(0.78)	(0.46)	(1.36)	(0.24)
AGE	.0312	.0315	.0327	0.326	.0401	.0436
	(3.28)	(3.01)	(3.42)	(3.10)	(3.45)	(3.50)
TEMPAN	.00964	.00765	.0156	.0235	.0206	.0274
	(1.28)	(0.81)	(1.74)	(1.93)	(1.49)	(1.55)
PRECAN	-.000154	-.0000892	-.00522	-.00551	-.00714	-.00363
	(0.00)	(0.00)	(1.40)	(1.21)	(1.28)	(0.47)
HUMIDAN	-.0131	-.0114	-.0119	.000376	-.0271	-.0186
	(1.66)	(1.31)	(1.42)	(0.00)	(1.72)	(1.11)
F or X2	44.05	44.57	48.90	43.15	49.06	54.62
N	8,483	6,886	8,486	6,795	5,780	5,007

positive in both equations, ozone was not significantly related to the likelihood of CRD in either specification of the model. Neither were the other air pollutants, with the exception of total suspended particulates in equation (31) where it had an unexpected sign. Nitrogen dioxide, which was sometimes positively and significantly related to acute respiratory illness among children in the analysis above, was positively but not significantly associated with CRD in equation (31).

The significant determinants of CRD in children were race (whites are more likely than blacks to have CRD) sex, and age. Since we are dealing with children in (30) and (31), the significance of the coefficient on the age variable is to be expected--the very youngest children would in general not have had sufficient time to develop or manifest CRD.

In equations (32) and (33), we replicate these results using air pollution averaged over as many of the six previous years as there were valid annual average air pollution data. For the very youngest children, the data in these two regressions may characterize concentrations at the monitors nearest their homes over their entire lifetime (if they have lived in the same location reported in the HIS interview). Using the multi-year averaged pollution data made little difference in the findings about ozone and CRD in children. In equation (32), however, sulfur dioxide entered as positively and significantly related to CRD. Race, sex, and age remained important explanatory variables under the alternative specifications.

Finally, we ran two regressions which included PGASPIPD, the probability that the child came from a household where natural gas was used as the cooking fuel. These are reported in equations (34) and (35). In neither were any of the air pollutants positively and significantly associated with CRD at the 5 percent level. In (35), however, carbon monoxide was significant at the 10 percent level, a finding for which we have no explanation.

In Equation (34), which includes three pollutants averaged over the period 1974-1979, the gas stove variable was positively associated with CRD in children at the 10 percent level. When carbon monoxide and nitrogen dioxide were included in (35), however, it was no longer significant. In (34) and (35) race, sex and age remained significantly associated with CRD in the children we examined.

To reiterate, we found no association between ozone and chronic respiratory disease in children. Admittedly, our search was a somewhat limited one, but only because our basic model--which often suggested some associations in our examinations of acute and chronic respiratory disease in adults--gave us no reason to believe such a search would be fruitful.

4.4 Cardiovascular and Other Chronic Diseases

To this point our analysis has been confined to acute morbidity due to respiratory and other causes, and also to chronic respiratory disease. However, air pollution may increase the likelihood of other diseases as well, including disease of the cardiovascular system.[15] Our final analysis of chronic illness concerns cardiovascular and other diseases.

Using a sample of 7,244 adults, of whom 733 (10.6 percent) reported having chronic cardiovascular disease (CCD), we estimated the models in equations (36) - (38) in Table 4-17. In none of the three models were any

Table 4-17. Analysis of Chronic Cardiovascular Disease

Model #	(36)	(37)	(38)
Dependent Var. Independent Var.	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)
INTERCEPT	-6.884 (6.12)	-7.713 (5.67)	-6.948 (6.17)
O3ANARAV	13.174 (1.61)	9.431 (0.98)	13.308 (1.62)
SPANGEAV	-.0000370 (0.00)	-.000575 (0.14)	-.0000693 (0.00)
S2ANARAV	.000741 (0.24)	.00148 (0.40)	.000635 (0.20)
COANRAV	-- --	.0532 (0.86)	-- --
N2ANARAV	-- --	-.000304 (0.10)	-- --
RACEW1BO	-.515 (4.26)	-.471 (3.49)	-.531 (4.38)
SEXM1FO	-.200 (2.19)	-.286 (2.84)	-.291 (3.07)
MARY1NO	.383 (3.98)	.326 (3.10)	.353 (3.65)
INCOMCON	-.0000293 (4.51)	-.0000216 (3.72)	-.0000239 (4.49)
SMOKY1NO	.132 (1.39)	.209 (2.01)	-- --
SCURRENT	-- --	-- --	.291 (2.76)
SFORMER	-- --	-- --	.396 (3.68)
FAT	1.404 (2.37)	1.444 (2.22)	1.418 (2.39)

Table 4-17 continued

Model #	(36)	(37)	(38)
Dependent Var. Independent Var.	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)	PRSCHRSP (LOGIT)
FATSQ	-.0852 (0.78)	-.0810 (0.67)	-.0869 (0.79)
AGE	-- --	-- --	-- --
AGESQ	.000549 (21.45)	.000564 (19.93)	.000547 (21.21)
EDCOMCON	-.0308 (2.32)	-.0214 (1.45)	-.0327 (2.46)
TEMPAN	.00485 (0.59)	.0122 (1.05)	.00480 (0.58)
PRECAN	.00585 (1.75)	.00587 (1.39)	.00590 (1.77)
HUMIDAN	.00862 (1.01)	.0101 (1.02)	.00852 (0.99)
F or X^2 N	949.12 7,244	773.78 6,008	962.52 7,244

of the air pollutants positively and significantly associated with CCD. A number of the other independent variables had the expected signs and were significant. For example, the dummy variable for race was highly significant and is consistent with earlier findings that blacks are more likely than whites to suffer from CCD. Income was negatively and significantly associated with CCD, as was education. Both findings confirm *a priori* expectations, as does the finding that the likelihood of CCD varied directly with weight. The square of age was by far the most significant explanatory variable. According to (36) - (38) women are more likely than men to report CCD, as are married as opposed to single individuals.

The crude dummy variable indicating current smoking status was significant at the 5 percent level in (37), although it was not in (36). As expected, it indicates that smoking is positively related to CCD. An additional regression identical to (36) was run substituting for SMOKY1N0, the pair of dummy variables distinguishing both current as well as former smokers. The results are reported in (38). Differentiating between those who never smoked and those who once smoked but no longer do makes a great difference. This model indicates that both current and former smokers have a significantly greater risk of CCD than non-smokers ($p < .01$). There is little change in the other signs where SCURRENT and SFORMER were included as possible determinants of CCD.

This analysis of CCD should be viewed as preliminary. Importantly, the regressions presented here contain no data on individuals' alcohol consumption (because no data were available from the HIS). This omission may be an important one in light of previous findings that CCD is positively and significantly related to alcohol.[16] If alcohol consumption does not vary systematically with air pollution, however, the coefficients on ozone and the other pollutants will not be biased.

Finally, based on the recommendations of a number of individuals we contacted at the outset of our study,[17] we explored in a preliminary way for possible associations between ozone and other air pollutants and several types of acute and chronic illness. The acute illnesses we explored included those of the eyes and ears as well as a group of infectious diseases. The chronic diseases included neoplasms, mental and nervous disorders, diseases of the digestive tract and several other illnesses lumped together under "other symptoms and ill-defined conditions" in the HIS diagnostic recode. Using our "basic model" we found no association between ozone and any of the acute or chronic illnesses we examined. These results are not presented here but are available from the authors.

Analysis of Multicollinearity

Multicollinearity in the linear model has been defined as the situation where "the rank of the regressor matrix X is less than K, the number of regressors," and "near multicollinearity" to be the case where typically "one of the regressors is 'almost' a linear combination of the others....in this case X^TX is 'almost' singular..."[18] The implications of multicollinearity and near multicollinearity for parameter estimation in the linear, nonlinear, and multinomial logit models are fairly well known and will not be reviewed here. Rather, in what follows we discuss the analyses of collinearity we have undertaken in the present context of assessing the effects of air pollution on human health.

First, we examine the question of correlation between the independent variables used in the models estimated above. Then, we describe analytical tools used to detect what is known as harmful or degrading collinearity.[19] We concentrate on collinearity in the linear model since most of the estimation done in this project uses ordinary least squares in the linear framework. Following this, we examine the empirical diagnostic results emanating from a subset of the models estimated, and assess the implications of these results for the interpretation of the coefficients estimated in this report.

The reasons underlying this exercise are several. Because the world is not orthogonal--that is, not all variables can be presumed to be independent of one another--the attempt to estimate partial derivatives will often be fraught with error. The diagnostics referred to above are invoked to indicate the extent to which our estimates are confounded by the presence of significant covariation between and among linear combinations of variables in the design matrix, and to identify what these covariates are. The diagnostics we use are particularly helpful in this respect. Whereas simple correlation coefficients allow one to see where the covariance between two variables might prove harmful to parameter estimation, this is not the case when the covariance arises due to a linear combination of more than two variables. For example, consider a linear model of the form

$$y = a + b_1X_1 + b_2X_2 + b_3X_3 + e$$

or

$$y = XB + e$$

Assume we are given the data

$$X = \begin{bmatrix} 1 & 2 & 2.1 & 0.1 \\ 1 & 7 & 7.1 & 0.1 \\ 1 & 14 & 14.0 & 0.1 \\ 1 & 3 & 2.9 & 0.1 \\ 1 & 6 & 0.1 & 5.9 \\ 1 & 5 & 0.1 & 5.1 \\ 1 & 2 & 0.1 & 2.1 \\ 1 & 6 & 0.1 & 6.1 \end{bmatrix}$$

The simple correlation matrix is

$$R = \begin{bmatrix} 1.00 & .83 & -.08 \\ & 1.00 & -.62 \\ & & 1.00 \end{bmatrix}$$

where $r_{ij} = s_{ij}/(s_i s_j)$, $i,j \in \{1, 2, 3\}$, the s_i and s_{ij} having the usual definitions as sample standard deviations and covariances.

Upon examination, one might be convinced that X_1 and X_2 are closely correlated ($r_{12} = .83$), but that the correlation between X_1 and X_3 is not very strong ($r_{13} = -.08$). However, consider $X_4 = X_2 + X_3$. It turns out that $r_{14} = .999$, i.e. X_1 is almost a perfect linear combination of X_2 and X_3. Due to this nearly perfect collinearity, parameter estimates will likely be seriously degraded in estimation of models such as that described above.

There is another rationale for exploring these issues of multicollinearity. There has been considerable concern expressed in previous epidemiological studies about the problems in interpreting coefficient estimates on individual pollution variables when such variables are correlated with each other, with meteorological variables, and with other explanatory variables. Even in work where much effort is devoted to examining multicollinearity among pollution variables, concern is often chiefly with simple correlations between two variables. Inasmuch as the primary objective of such studies has been to assess the influence of each explanatory variable individually on health status, such concern is rightly placed.

The bulk of the interest in collinearity problems in statistical studies like the present one has focused on the covariance of the pollution measures employed and on the related point of covariances between and among pollution summary measures and meteorological variables. In areas where some ambient pollution levels are high, others tend to be high also. For instance, because ozone and other photochemical oxidants result in part from vehicular emissions of hydrocarbons and nitrogen oxides, ozone tends to be correlated with carbon monoxide, another mobile source pollutant. Similarly, the nature of the physical world is such that when certain meteorological conditions prevail, so too do levels of some pollutants. For instance, sunlight and heat tend to "cook" hydrocarbons and oxides of nitrogen to form atmospheric ozone. Thus, ozone and temperature are positively correlated.

We have earlier examined the bivariate correlations between the explanatory variables used in our analysis. Recall that it was shown that certain of varibles tend to be almost perfectly correlated (e.g. $r_{AGE,AGESQ}=.982$), other pairs quite significantly but not so perfectly correlated (e.g. $r_{O3NR01,AVMAXTMP}=.533$, while other variables tended to be virtually orthogonal (e.g. $r_{03NR01,AGE}=-.003$). We now turn our focus to more sophisticated collinearity diagnostics, reemphasizing that examination of the simple correlation coefficients can be quite misleading in assessing the potentially harmful effects of multicollinearity on parameter estimation.[20]

In the presence of severe collinearity, the least squares estimator of β, $\hat{\beta} = (X^TX)^{-1}X^Ty$ and the estimated covariance matrix, $s^2(X^TX)^{-1}$ tend to be numerically unstable. In both cases, this is due to the fact that X^TX tends towards singularity as the severity of the collinearity increases, with the determinant (X^TX) tending to vanish. Such collinearity corresponds to the existence of at least one small eigenvalue of the X^TX matrix, with the numerical-analytic concept of condition numbers or condition indexes employed to indicate just what constitutes a "small" eigenvalue. Simply put, large condition indexes are representative of small eigenvalues. It is important to note that in the linear model, the problem of collinearity is manifest solely in the X matrix without any reference to the error structure or dependent variables under consideration.

There is no statistical basis on which one can judge when harmful or degrading collinearity is present even given the magnitude of the condition indexes. Experimental experience has shown, however, that condition indexes on the order of 5 to 10 tend to be associated with weak linear dependencies among the columns of X, whereas indexes of 30 and above tend to be associated with moderate to strong (i.e. harmful and degrading) collinearity.

The effects of collinearity have also been examined in a more specific sense, that is, via the decomposition of the variance of each estimated parameter. In this approach, the estimated variance of each regression coefficient is decomposed into a sum of terms, each term associated with one singular value (the singular values are the positive square roots of the eigenvalues). For each component $\hat{\beta}_k$ of $\hat{\beta}$, it has been shown that $\mathrm{var}(\hat{\beta}_k) = \sigma^2 \sum_{j=1}^{p} (v_{kj}^2 / \mu_j^2)$. Here, $[v_{kj}]$ are the typical elements of V where V results from the singular value decomposition of X, $X = UDV^T$, where it can be shown that $\mathrm{var}(\hat{\beta}) = \sigma^2 (X^T X)^{-1} = \sigma^2 V D^{-2} V^T$. μ_j is the j-th singular value. If one defines $\phi_{kj} = (v_{kj}^2 / \mu_j^2)$ and $\phi_k = \sum_{j=1}^{p} \phi_{kj}$, then the k,j-th variance decomposition proportion -- i.e. the percentage of var (β_k) associated with the j-th component of its decomposition -- is calculated as $\Pi_{jk} = (\phi_{kj} / \phi_k)$. (These collinearity diagnostics can be calculated in the case of the linear model in SAS's PROC REG specifying the COLLIN option in the MODEL statement.)

Strictly speaking, the analytics described above are not fully appropriate in the case of the binomial logit model.[21] In the linear model, the problem of collinearity was shown above to be fully attributable to characteristics of the X matrix, specifically, significant linear relationships among the columns of X. Here, the covariance matrix of the estimated parameters is merely a scalar multiple of the $(X^T X)^{-1}$ matrix. In the binomial -- or, more generally, multinomial -- logit model, the estimated covariance matrix equals J^{-1}, where J, the information matrix, depends not only on the data matrix X but also on the estimated logistic distribution function $F(X\beta)$. It has been demonstrated that the fundamentals of the regression diagnostics for the linear case can be applied to the analysis of collinearity in the logit case, although one must work with the information matrix instead of the $(X^TX)^{-1}$ matrix.

Although we have almost completed work on the software capable of performing the collinearity diagnostics in the logit case, we have not yet been able to perform the necessary benchmark tests. Because logistic estimation has been important in many facets of this study, however, we have performed collinearity diagnostics on the X matrixes used in the binomial logit models *as if* such models were estimated by OLS with binary dependent variables. Although different results would be forthcoming were the estimated logit covariance matrixes to be decomposed, the present exercise is nonetheless illustrative. Of course, for the models estimated by OLS, the collinearity diagnostics described above are entirely appropriate.

We now turn to a discusssion of the collinearity diagnostics as applied to twelve of the models estimated in our project and discussed earlier in this chapter. This is followed by a tabulation of the full set of collinearity diagnostics produced by SAS: eigenvalues, condition indexes, and the variance proportions matrix.

The most important result in the context of the present analysis is that the ozone variable--however summarized in the models under consideration--cannot be said to be a partner with any other variable or set of variables in harmful collinearity. The criterion on which this conclusion is based is the following: given an associated condition index of 30 or greater, the proportions of the variances of the ozone parameter and that of at least one other variable not be greater than 0.50. The single exception to this occurs in Acute Model 26 where the ozone measure and the ozone-temperature interactive term measure appear to be involved in some potentially confounding collinearity. This would be expected given the nature of the variates.

Equally noteworthy is that, by the same criterion, all pollution measures in each of the twelve models (except Acute Model 26) withstand scrutiny for involvement in harmful collinearity, as do all the climatological measures. Of course, the criterion established here is an *ad hoc* one for the reasons discussed earlier. However, it is a reasonable and often-utilized one.

What would by our standards be termed harmful collinearity exists where one would expect it, i.e., among functions of the same variable. In many of the models estimated over the course of this study, this refers to the (AGE, AGESQ) and (FAT, FATSQ) pairs. The inclusion of both linear and quadratic terms of the same variable is often found to be a major contributor to collinearity. This is the case in the present analysis. However, *a priori* restrictions on the specification (i.e. "leaving out" either the linear or quadratic term) constitutes the introduction of a prior restriction on the estimating equation, which can itself have undesirable properties. Testing for both linear and quadratic effects in the cases of AGE and FAT seems quite plausible. Therefore, one is left with a tradeoff whose resolution is beyond the scope of the present effort. We have little to say about the involvement of other variates in harmful collinearity. It is generally true that the intercept is involved in the collinearity with FAT and FATSQ, suggesting a strong relationship of the form FAT = c + b*FATSQ (plus some error). Otherwise, there appears to be little degrading collinearity present in the models scrutinized here. In further research, we hope to analyze the logit models in a more sophisticated manner, and will perhaps have different conclusions to draw after such an examination.

Model: Chronic 25

```
COLLINEARITY DIAGNOSTICS          VARIANCE PROPORTIONS

                      CONDITION   PORTION  PORTION  PORTION  PORTION  PORTION  PORTION  PORTION  PORTION  PORTION  PORTION  PORTION  PORTION  PORTION  PORTION  PORT
NUMBER EIGENVALUE         INDEX  INTERCEP BJAMARAV BPANGEAV BZAMARAV RACEWIBO  BEXMIPO  MARVINO INCOMCON      FAT    FAT80    AGE80 EDCOMCON   TEMPAN   PRECAN HUMI

     1     12.804         1.000    0.0000   0.0003   0.0002   0.0006   0.0007   0.0013   0.0014   0.0009   0.0000   0.0000   0.0010   0.0003   0.0000   0.0004  0.0
     2   0.499402         5.064    0.0000   0.0006   0.0005   0.0104   0.0014   0.6615   0.0063   0.0001   0.0000   0.0001   0.0101   0.0000   0.0001   0.0018  0.0
     3   0.422002         5.508    0.0000   0.0000   0.0007   0.0163   0.0000   0.0178   0.6097   0.0146   0.0000   0.0001   0.0726   0.0011   0.0001   0.0032  0.0
     4   0.352020         6.031    0.0000   0.0001   0.0000   0.0000   0.0022   0.0127   0.0493   0.0836   0.0000   0.0000   0.7059   0.0024   0.0000   0.0000  0.0
     5   0.221226         7.608    0.0000   0.0241   0.0030   0.3750   0.0436   0.0119   0.0609   0.0003   0.0000   0.0000   0.0004   0.0018   0.0014   0.0290  0.0
     6   0.187317         8.268    0.0000   0.0113   0.0075   0.0006   0.0197   0.0137   0.2170   0.5308   0.0001   0.0010   0.1455   0.0022   0.0003   0.0043  0.0
     7   0.170675         8.662    0.0000   0.0101   0.0299   0.1333   0.0147   0.0260   0.0006   0.0042   0.0001   0.0010   0.0031   0.0002   0.0001   0.2414  0.0
     8   0.125540        10.099    0.0000   0.0018   0.0027   0.0015   0.4851   0.0653   0.0277   0.2159   0.0002   0.0030   0.0190   0.0000   0.0000   0.0481  0.0
     9   0.103991        11.096    0.0002   0.0353   0.0108   0.0140   0.3663   0.0847   0.0075   0.0016   0.0003   0.0000   0.0000   0.0082   0.0034   0.0515  0.0
    10   0.049034        16.160    0.0001   0.1907   0.0050   0.0066   0.0154   0.0018   0.0164   0.1355   0.0000   0.0000   0.0120   0.4539   0.0000   0.1007  0.0
    11   0.033794        19.465    0.0005   0.6870   0.3376   0.0855   0.0001   0.0024   0.0004   0.0108   0.0000   0.0001   0.0112   0.1150   0.0139   0.0066  0.0
    12   0.018716        26.156    0.0070   0.0304   0.4734   0.0349   0.0216   0.0006   0.0003   0.0001   0.0000   0.0010   0.0132   0.1832   0.0449   0.4714  0.1
    13   0.008324        39.221    0.0001   0.0017   0.1241   0.2259   0.0100   0.0005   0.0013   0.0001   0.0000   0.0000   0.0019   0.0000   0.5514   0.0095  0.3
    14   0.002318        74.316    0.2699   0.0057   0.0042   0.0966   0.0146   0.0050   0.0002   0.0014   0.0330   0.0401   0.0017   0.0299   0.3751   0.0308  0.5
    15 .000372384       185.432    0.7221   0.0005   0.0004   0.0007   0.0045   0.0960   0.0012   0.0000   0.9651   0.9439   0.0018   0.0006   0.0092   0.0012  0.0
```

Model: Chronic 23

OLLINEARITY DIAGNOSTICS VARIANCE PROPORTIONS

UMBER	EIGENVALUE	CONDITION INDEX	PORTION INTERCEP	PORTION O3ANARAV	PORTION SPANGEAV	PORTION S2ANARAV	PORTION RACENIBO	PORTION SEXMIFO	PORTION MARYING	PORTION INCOMCON	PORTION FAT	PORTION FAT80	PORTION AGE80	PORTION EDCOMCON	PORTION TEMPAN	PORTION PRECAN	PORTI HUMI
1	12.485	1.000	0.0000	0.0003	0.0002	0.0007	0.0007	0.0013	0.0016	0.0009	0.0000	0.0000	0.0013	0.0004	0.0000	0.0005	0.0
2	0.692900	4.245	0.0000	0.0001	0.0001	0.0007	0.0004	0.5670	0.0100	0.0039	0.0000	0.0000	0.1614	0.0002	0.0000	0.0004	0.0
3	0.441121	5.320	0.0000	0.0000	0.0001	0.0049	0.0006	0.0554	0.4633	0.0564	0.0000	0.0001	0.2024	0.0013	0.0000	0.0009	0.0
4	0.411776	5.506	0.0000	0.0013	0.0012	0.0142	0.0017	0.1764	0.3369	0.0130	0.0000	0.0000	0.2088	0.0055	0.0002	0.0016	0.0
5	0.230872	7.230	0.0000	0.0197	0.0034	0.3491	0.0344	0.0013	0.0502	0.0082	0.0000	0.0000	0.0125	0.0015	0.0010	0.0395	0.0
6	0.187122	8.168	0.0000	0.0173	0.0200	0.0206	0.0123	0.0049	0.0723	0.3334	0.0001	0.0010	0.1440	0.0021	0.0082	0.0923	0.0
7	0.179362	8.343	0.0000	0.0009	0.0002	0.1569	0.0551	0.0438	0.0008	0.0511	0.0002	0.0031	0.0011	0.0000	0.0003	0.1275	0.0
8	0.131412	9.747	0.0001	0.0012	0.0002	0.0040	0.1311	0.0513	0.0559	0.4553	0.0002	0.0050	0.0443	0.0104	0.0007	0.0075	0.0
9	0.109072	10.699	0.0001	0.0345	0.0132	0.0049	0.6761	0.0100	0.0030	0.0098	0.0001	0.0036	0.0344	0.0174	0.0020	0.0007	0.0
10	0.060651	14.347	0.0000	0.1092	0.0062	0.0012	0.0096	0.0061	0.0004	0.0610	0.0000	0.0003	0.0526	0.7599	0.0005	0.0033	0.0
11	0.032625	19.562	0.0010	0.0056	0.2352	0.0695	0.0023	0.0024	0.0001	0.0016	0.0000	0.0007	0.0103	0.0902	0.0197	0.0002	0.0
12	0.018836	25.745	0.0066	0.0037	0.5591	0.0272	0.0425	0.0000	0.0022	0.0012	0.0004	0.0022	0.0126	0.0094	0.0325	0.5092	0.1
13	0.008210	38.976	0.0001	0.0016	0.1513	0.2051	0.0052	0.0009	0.0003	0.0001	0.0000	0.0000	0.0001	0.0008	0.4960	0.0112	0.0
14	0.002581	69.553	0.2490	0.0027	0.0000	0.1393	0.0166	0.0001	0.0001	0.0040	0.0350	0.0411	0.0080	0.0187	0.4307	0.0211	0.4
15	.000399295	176.826	0.7429	0.0020	0.0007	0.0017	0.0074	0.0711	0.0081	0.0001	0.9639	0.9427	0.0062	0.0024	0.0155	0.0063	0.0

Model: Chronic 24

COLLINEARITY DIAGNOSTICS VARIANCE PROPORTIONS

NUMBER	EIGENVALUE	CONDITION INDEX	PORTION INTERCEP	PORTION O3ANARAV	PORTION SPANGEAV	PORTION S2ANARAV	PORTION RACENIBO	PORTION SEXMIFO	PORTION MARYING	PORTION INCOMCON	PORTION FAT	PORTION FAT80	PORTION AGE80	PORTION EDCOMCON	PORTION TEMPAN	PORTION PRECAN	PORT HUMI
1	13.210	1.000	0.0000	0.0003	0.0002	0.0006	0.0004	0.0011	0.0010	0.0007	0.0000	0.0000	0.0010	0.0003	0.0000	0.0004	0.0
2	0.394061	5.790	0.0000	0.0003	0.0001	0.0069	0.0013	0.5336	0.0004	0.0208	0.0000	0.0001	0.0607	0.0049	0.0001	0.0062	0.0
3	0.329086	6.336	0.0000	0.0000	0.0003	0.0276	0.0000	0.1063	0.0628	0.0887	0.0000	0.0000	0.3260	0.0033	0.0000	0.0037	0.0
4	0.242985	7.373	0.0000	0.0000	0.0006	0.1086	0.0000	0.0704	0.4606	0.0025	0.0000	0.0000	0.1715	0.0034	0.0000	0.0034	0.0
5	0.218663	7.773	0.0000	0.0392	0.0242	0.0971	0.0013	0.0236	0.0881	0.0034	0.0000	0.0001	0.0010	0.0015	0.0014	0.1219	0.0
6	0.167768	8.874	0.0000	0.0000	0.0150	0.3193	0.0000	0.0018	0.1630	0.0007	0.0001	0.0000	0.0017	0.0016	0.0007	0.1379	0.0
7	0.149314	9.406	0.0000	0.0012	0.0009	0.0034	0.0217	0.0385	0.1164	0.2702	0.0003	0.0064	0.2140	0.0146	0.0000	0.0031	0.0
8	0.104378	11.250	0.0002	0.0221	0.0021	0.0242	0.0347	0.0832	0.0786	0.3561	0.0003	0.0097	0.1131	0.0062	0.0020	0.0549	0.0
9	0.072652	13.484	0.0000	0.0938	0.0073	0.0017	0.7304	0.0122	0.0011	0.0489	0.0000	0.0016	0.0103	0.0001	0.0010	0.0421	0.0
10	0.046692	16.820	0.0000	0.1662	0.0000	0.0061	0.1081	0.0007	0.0226	0.1699	0.0000	0.0002	0.0467	0.7495	0.0000	0.0411	0.0
11	0.034017	19.706	0.0008	0.6373	0.3078	0.0745	0.0187	0.0020	0.0022	0.0247	0.0000	0.0011	0.0140	0.1260	0.0179	0.0252	0.0
12	0.018685	26.590	0.0079	0.0327	0.4472	0.0143	0.0609	0.0002	0.0001	0.0003	0.0006	0.0029	0.0227	0.0718	0.0366	0.4898	0.1
13	0.008343	39.792	0.0000	0.0034	0.1810	0.2276	0.0010	0.0026	0.0021	0.0010	0.0000	0.0000	0.0019	0.0002	0.5710	0.0319	0.3
14	0.002648	70.634	0.2859	0.0027	0.0081	0.0080	0.0149	0.0011	0.0005	0.0069	0.0411	0.0492	0.0125	0.0112	0.3668	0.0300	0.4
15	.000484122	165.188	0.7050	0.0007	0.0051	0.0002	0.0059	0.1145	0.0005	0.0055	0.9574	0.9277	0.0027	0.0054	0.0024	0.0084	0.0

Model: Chronic 20

```
COLLINEARITY DIAGNOSTICS          VARIANCE PROPORTIONS

                     CONDITION    PORTION  PORTION  PORTION  PORTION  PORTION  PORTION  PORTION  PORTION  PORTION  PORTION  PORTION
NUMBER EIGENVALUE        INDEX   INTERCEP O3ANARAV SPANGEAV SZANARAV RACEWIND  SEXMIFD  MARVIND INCOMCON      FAT    FATSQ    AGESQ

     1     13.302        1.000     0.0000   0.0002   0.0001   0.0004   0.0006   0.0013   0.0013   0.0008   0.0000   0.0000   0.0012
     2   0.779196        4.132     0.0000   0.0003   0.0001   0.0004   0.0004   0.0006   0.0011   0.0002   0.0000   0.0000   0.0203
     3   0.733416        4.259     0.0000   0.0000   0.0000   0.0001   0.0002   0.0193   0.0024   0.0081   0.0000   0.0000   0.0021
     4   0.536828        4.978     0.0000   0.0003   0.0002   0.0034   0.0015   0.6776   0.0163   0.0019   0.0000   0.0000   0.0354
     5   0.400370        5.764     0.0000   0.0001   0.0001   0.0074   0.0019   0.0679   0.6043   0.0519   0.0000   0.0001   0.0142
     6   0.388174        5.854     0.0000   0.0012   0.0003   0.0009   0.0012   0.0240   0.0579   0.0431   0.0000   0.0000   0.5912
     7   0.246784        7.342     0.0000   0.0357   0.0063   0.0515   0.0163   0.0000   0.0212   0.0010   0.0000   0.0000   0.0081
     8   0.190119        8.365     0.0000   0.0047   0.0027   0.0231   0.1093   0.0216   0.2006   0.3160   0.0001   0.0013   0.1641
     9   0.131208       10.069     0.0000   0.0000   0.0005   0.0614   0.2645   0.0696   0.0699   0.2568   0.0003   0.0057   0.0376
    10   0.116335       10.693     0.0000   0.0063   0.0037   0.0929   0.4600   0.0395   0.0004   0.1021   0.0002   0.0049   0.0649
    11   0.081793       12.753     0.0002   0.0077   0.0001   0.4121   0.0917   0.0044   0.0180   0.1512   0.0000   0.0015   0.0000
    12   0.052436       15.928     0.0001   0.0534   0.0022   0.0093   0.0000   0.0078   0.0013   0.0629   0.0000   0.0000   0.0369
    13   0.022183       24.488     0.0023   0.8161   0.1643   0.1168   0.0039   0.0002   0.0007   0.0031   0.0001   0.0008   0.0048
    14   0.011972       33.334     0.0071   0.0001   0.4010   0.0017   0.0215   0.0000   0.0000   0.0053   0.0005   0.0018   0.0061
    15   0.003726       59.748     0.0006   0.0719   0.3929   0.1954   0.0077   0.0006   0.0007   0.0011   0.0001   0.0001   0.0000
    16   0.002676       70.502     0.2656   0.0016   0.0222   0.0227   0.0161   0.0002   0.0000   0.0024   0.0292   0.0340   0.0089
    17 .000389212      184.872     0.7238   0.0003   0.0031   0.0003   0.0032   0.0652   0.0040   0.0001   0.9695   0.9483   0.0021

        PORTION  PORTION  PORTION  PORTION  PORTION  PORTION
NUMBER ACIDSOYN ENCOMCON   TEMPAN   PRECAN  HUMIDAN POLLENAN

     1   0.0011   0.0003   0.0000   0.0002   0.0000   0.0010
     2   0.4646   0.0001   0.0006   0.0002   0.0000   0.3239
     3   0.4337   0.0000   0.0008   0.0004   0.0000   0.4223
     4   0.0265   0.0001   0.0001   0.0007   0.0001   0.0102
     5   0.0065   0.0000   0.0000   0.0016   0.0001   0.0099
     6   0.0412   0.0078   0.0001   0.0002   0.0001   0.0075
     7   0.0025   0.0000   0.0001   0.0679   0.0000   0.0406
     8   0.0099   0.0037   0.0001   0.0034   0.0001   0.0008
     9   0.0121   0.0010   0.0030   0.0009   0.0001   0.0033
    10   0.0059   0.0072   0.0001   0.0048   0.0001   0.0074
    11   0.0001   0.0004   0.0027   0.1035   0.0007   0.0619
    12   0.0033   0.7933   0.0026   0.0079   0.0006   0.0096
    13   0.0000   0.0454   0.0102   0.1465   0.0161   0.0008
    14   0.0003   0.0305   0.0009   0.3010   0.2515   0.0012
    15   0.0004   0.0004   0.0224   0.3224   0.3057   0.0968
    16   0.0004   0.0280   0.1590   0.0532   0.4167   0.0065
    17   0.0006   0.0818   0.8015   0.0062   0.0081   0.0001
```

MODEL: CHRONIC 15

```
COLLINEARITY DIAGNOSTICS          VARIANCE PROPORTIONS

                      CONDITION    PORTION   PORTION   PORTION   PORTION   PORTION   PORTION   PORTION   PORTION   PORTION   PORTION   PORTION
NUMBER EIGENVALUE         INDEX   INTERCEP  PRASPIPD  D3ANARAV  SPANGEAV  S2ANARAV  RACEWIRO   SEXMIFO   MARYING  INCOMCON       FAT     FATSQ

     1     13.916         1.000     0.0000    0.0002    0.0002    0.0001    0.0003    0.0006    0.0011    0.0012    0.0007    0.0000    0.0000
     2   0.757893         4.205     0.0000    0.0002    0.0001    0.0000    0.0001    0.0001    0.0093    0.0036    0.0001    0.0000    0.0000
     3   0.533621         5.107     0.0000    0.0014    0.0003    0.0002    0.0027    0.0007    0.6905    0.0003    0.0018    0.0000    0.0001
     4   0.412238         5.810     0.0000    0.0047    0.0006    0.0006    0.0075    0.0003    0.0095    0.5312    0.0014    0.0000    0.0000
     5   0.392477         5.954     0.0000    0.0001    0.0001    0.0000    0.0007    0.0015    0.0514    0.1234    0.0913    0.0000    0.0001
     6   0.240453         7.607     0.0000    0.0044    0.0308    0.0087    0.0697    0.0401    0.0032    0.0582    0.0003    0.0000    0.0000
     7   0.197796         8.388     0.0000    0.0262    0.0149    0.0099    0.0244    0.0518    0.0001    0.1961    0.2310    0.0000    0.0001
     8   0.161024         9.296     0.0000    0.0254    0.0001    0.0011    0.0550    0.0139    0.0765    0.0040    0.1424    0.0004    0.0056
     9   0.125795        10.518     0.0000    0.0040    0.0007    0.0004    0.0006    0.4496    0.0326    0.0507    0.3960    0.0001    0.0023
    10   0.096864        11.986     0.0002    0.0004    0.0187    0.0016    0.0472    0.3778    0.0407    0.0146    0.0269    0.0002    0.0070
    11   0.062520        14.919     0.0000    0.4899    0.0032    0.0448    0.1703    0.0214    0.0001    0.0002    0.0018    0.0000    0.0000
    12   0.052010        16.233     0.0000    0.0609    0.1047    0.0192    0.0044    0.0001    0.0040    0.0055    0.1027    0.0000    0.0005
    13   0.028551        22.077     0.0016    0.0211    0.0144    0.1727    0.0471    0.0037    0.0005    0.0006    0.0016    0.0001    0.0010
    14   0.016392        29.136     0.0051    0.0409    0.0026    0.6260    0.1038    0.0267    0.0001    0.0000    0.0002    0.0005    0.0016
    15   0.003920        59.580     0.0064    0.2467    0.0063    0.0685    0.4632    0.0037    0.0003    0.0000    0.0007    0.0008    0.0013
    16   0.001650        91.824     0.2908    0.0713    0.0023    0.0391    0.0000    0.0046    0.0034    0.0000    0.0011    0.0467    0.0704
    17 .000406644       184.988     0.6956    0.0021    0.0000    0.0070    0.0000    0.0040    0.0706    0.0019    0.0000    0.0313    0.9099

         PORTION   PORTION   PORTION    PORTION    PORTION    PORTION
NUMBER     AGESQ  ACIGSOYN  EDCOMCON     TEMPAN     PRECAN    MINIDAN

     1    0.0011    0.0016    0.0003     0.0000     0.0003     0.0000
     2    0.0174    0.0007    0.0000     0.0000     0.0000     0.0000
     3    0.0164    0.0356    0.0001     0.0000     0.0011     0.0000
     4    0.1593    0.0349    0.0021     0.0001     0.0014     0.0000
     5    0.4695    0.0121    0.0044     0.0000     0.0001     0.0000
     6    0.0063    0.0001    0.0014     0.0006     0.0745     0.0001
     7    0.0952    0.0000    0.0069     0.0000     0.0436     0.0000
     8    0.0837    0.0007    0.0000     0.0004     0.0366     0.0002
     9    0.0750    0.0162    0.0011     0.0000     0.0274     0.0001
    10    0.0089    0.0009    0.0496     0.0026     0.0694     0.0008
    11    0.0058    0.0003    0.0625     0.0018     0.0181     0.0008
    12    0.0289    0.0055    0.6698     0.0003     0.1086     0.0003
    13    0.0135    0.0022    0.1088     0.0142     0.0273     0.0064
    14    0.0096    0.0002    0.0706     0.0204     0.4217     0.0593
    15    0.0008    0.0005    0.0016     0.0455     0.0206     0.2567
    16    0.0652    0.0005    0.0171     0.1056     C.1318     0.6403
    17    0.0036    0.0006    0.0048     0.0091     0.0202     0.0369
```

Model: Chronic 10

```
COLLINEARITY DIAGNOSTICS          VARIANCE PROPORTIONS

                      CONDITION   PORTION  PORTION  PORTION  PORTION  PORTION  PORTION  PORTION  PORTION  PORTION  PORTION  PORTION
NUMBER EIGENVALUE         INDEX  INTERCEP O3ANARAV SPANGEAV SZANARAV COANARAV NZANARAV RACEWIBO  SEXMIFO  MARYINO INCOMCON      FAT

     1     14.801         1.000    0.0000   0.0002   0.0001   0.0004   0.0002   0.0001   0.0005   0.0010   0.0010   0.0006   0.0000
     2   0.758638         4.417    0.0000   0.0001   0.0001   0.0002   0.0002   0.0001   0.0001   0.0189   0.0014   0.0001   0.0000
     3   0.534843         5.261    0.0000   0.0003   0.0003   0.0036   0.0010   0.0006   0.0005   0.6723   0.0130   0.0020   0.0000
     4   0.415256         5.970    0.0000   0.0007   0.0009   0.0066   0.0053   0.0028   0.0002   0.0107   0.4230   0.0000   0.0000
     5   0.390766         6.154    0.0000   0.0000   0.0000   0.0016   0.0003   0.0001   0.0011   0.0545   0.2201   0.0952   0.0000
     6   0.276545         7.316    0.0000   0.0140   0.0090   0.0451   0.0057   0.0172   0.0013   0.0002   0.0113   0.0101   0.0000
     7   0.203710         8.524    0.0000   0.0033   0.0003   0.1487   0.0310   0.0080   0.0590   0.0086   0.2524   0.1429   0.0000
     8   0.170549         9.316    0.0001   0.0063   0.0007   0.0411   0.0167   0.0031   0.0102   0.0580   0.0267   0.2720   0.0004
     9   0.122208        11.002    0.0001   0.0020   0.0008   0.0012   0.0019   0.0026   0.3542   0.0524   0.0352   0.3534   0.0002
    10   0.099127        12.219    0.0001   0.0037   0.0005   0.2179   0.0440   0.0021   0.3490   0.0165   0.0078   0.0054   0.0001
    11   0.088049        12.965    0.0001   0.0698   0.0272   0.1250   0.2531   0.0074   0.1618   0.0174   0.0001   0.0002   0.0000
    12   0.053359        16.655    0.0000   0.1275   0.0065   0.0034   0.0002   0.0002   0.0050   0.0030   0.0034   0.1060   0.0000
    13   0.029069        22.260    0.0002   0.2318   0.0222   0.0139   0.4431   0.4776   0.0002   0.0033   0.0002   0.0007   0.0000
    14   0.028345        22.851    0.0013   0.5261   0.4110   0.0185   0.0133   0.1080   0.0043   0.0008   0.0005   0.0064   0.0001
    15   0.017774        28.857    0.0078   0.0005   0.3910   0.0584   0.1006   0.1485   0.0334   0.0000   0.0007   0.0009   0.0005
    16   0.007599        44.133    0.0011   0.0002   0.1129   0.1572   0.0162   0.0000   0.0029   0.0002   0.0001   0.0000   0.0002
    17   0.001746        92.064    0.2972   0.0117   0.0141   0.1540   0.0636   0.2118   0.0118   0.0013   0.0002   0.0032   0.0588
    18 .000414754       188.909    0.6928   0.0018   0.0032   0.0034   0.0035   0.0091   0.0046   0.0810   0.0028   0.0000   0.9398

        PORTION   PORTION  PORTION  PORTION  PORTION  PORTION  PORTION
NUMBER    FATSQ     AGESQ NCIGSDYN EDCOMCON   TEMPAN   PRECAN  HUMIDAN

     1   0.0000    0.0010   0.0009   0.0002   0.0000   0.0003   0.0000
     2   0.0000    0.0164   0.8675   0.0000   0.0000   0.0000   0.0000
     3   0.0001    0.0093   0.0627   0.0001   0.0000   0.0049   0.0000
     4   0.0000    0.2552   0.0286   0.0022   0.0000   0.0000   0.0001
     5   0.0001    0.3738   0.0072   0.0034   0.0001   0.0001   0.0001
     6   0.0001    0.0007   0.0052   0.0007   0.0000   0.0001   0.0000
     7   0.0000    0.0503   0.0000   0.0086   0.0001   0.1151   0.0001
     8   0.0049    0.1529   0.0021   0.0002   0.0003   0.0000   0.0001
     9   0.0045    0.0506   0.0159   0.0050   0.0004   0.0038   0.0003
    10   0.0031    0.0081   0.0000   0.0224   0.0002   0.0250   0.0004
    11   0.0021    0.0070   0.0004   0.0164   0.0016   0.1150   0.0006
    12   0.0003    0.0357   0.0033   0.7291   0.0004   0.0099   0.0009
    13   0.0002    0.0053   0.0016   0.0561   0.0001   0.1070   0.0000
    14   0.0008    0.0053   0.0017   0.0539   0.0065   0.0147   0.0159
    15   0.0024    0.0139   0.0001   0.0020   0.0102   0.0084   0.0014
    16   0.0002    0.0002   0.0016   0.0002   0.0244   0.5798   0.0860
    17   0.0653    0.0101   0.0008   0.0173   0.3187   0.0002   0.4357
    18   0.9160    0.0043   0.0004   0.0023   0.6127   0.0156   0.4310
                                              0.0244   0.0042   0.0275
```

Model: Chronic 9

```
COLLINEARITY DIAGNOSTICS          VARIANCE PROPORTIONS

                     CONDITION   PORTION  PORTION  PORTION  PORTION  PORTION  PORTION  PORTION  PORTION  PORTION  PORTION  PORTION
NUMBER EIGENVALUE        INDEX  INTERCEP O3AMARAV SPANOEAV SZANAMAV RACEW180  SEXMIFO  MARYINO INCOMCON      FAT    FATSQ    AGESQ

     1     13.010        1.000    0.0000   0.0003   0.0002   0.0006   0.0006   0.0013   0.0013   0.0008   0.0000   0.0000   0.0012
     2   0.751712        4.160    0.0000   0.0001   0.0001   0.0003   0.0001   0.0148   0.0023   0.0001   0.0000   0.0000   0.0207
     3   0.526920        4.969    0.0000   0.0003   0.0003   0.0044   0.0009   0.6930   0.0008   0.0014   0.0000   0.0001   0.0255
     4   0.397156        5.724    0.0000   0.0006   0.0003   0.0028   0.0011   0.0360   0.0112   0.0583   0.0000   0.0000   0.5994
     5   0.384248        5.789    0.0000   0.0005   0.0007   0.0216   0.0001   0.0142   0.7040   0.0333   0.0000   0.0000   0.0019
     6   0.226212        7.584    0.0000   0.0258   0.0056   0.3408   0.0245   0.0064   0.0633   0.0012   0.0000   0.0000   0.0042
     7   0.181252        8.472    0.0000   0.0157   0.0265   0.0658   0.0091   0.0020   0.1243   0.2810   0.0000   0.0004   0.1133
     8   0.174553        8.633    0.0000   0.0002   0.0068   0.1148   0.0398   0.0360   0.0234   0.1386   0.0002   0.0035   0.0943
     9   0.123488       10.264    0.0001   0.0018   0.0001   0.0042   0.1274   0.0768   0.0573   0.3614   0.0003   0.0066   0.0448
    10   0.101479       11.323    0.0001   0.0394   0.0121   0.0070   0.6642   0.0272   0.0048   0.0076   0.0001   0.0043   0.0188
    11   0.054799       15.409    0.0000   0.1370   0.0047   0.0029   0.0193   0.0015   0.0032   0.1091   0.0000   0.0001   0.0378
    12   0.033535       19.697    0.0008   0.7549   0.2791   0.0774   0.0015   0.0003   0.0000   0.0056   0.0000   0.0005   0.0119
    13   0.018992       26.173    0.0072   0.0167   0.5150   0.0263   0.0359   0.0004   0.0008   0.0004   0.0005   0.0023   0.0147
    14   0.008326       39.525    0.0000   0.0016   0.1466   0.2169   0.0054   0.0000   0.0002   0.0000   0.0000   0.0000   0.0005
    15   0.002517       71.901    0.2676   0.0036   0.0026   0.1136   0.0148   0.0011   0.0000   0.0010   0.0375   0.0444   0.0089
    16 .000422842      175.411    0.7241   0.0013   0.0013   0.0009   0.0054   0.0881   0.0030   0.0002   0.9613   0.9379   0.0081


        PORTION  PORTION  PORTION  PORTION  PORTION
NUMBER NCIGSDYN EDCOMCON   TEMPAN   PRECAN  HUMIDAN

     1   0.0012   0.0003   0.0000   0.0004   0.0000
     2   0.8707   0.0001   0.0000   0.0001   0.0000
     3   0.0624   0.0001   0.0001   0.0016   0.0001
     4   0.0341   0.0070   0.0001   0.0004   0.0001
     5   0.0037   0.0001   0.0001   0.0027   0.0001
     6   0.0003   0.0016   0.0013   0.0483   0.0000
     7   0.0003   0.0031   0.0001   0.1068   0.0000
     8   0.0014   0.0004   0.0003   0.1139   0.0002
     9   0.0174   0.0048   0.0006   0.0725   0.0006
    10   0.0008   0.0106   0.0024   0.0211   0.0005
    11   0.0030   0.7425   0.0001   0.0056   0.0000
    12   0.0017   0.0987   0.0174   0.0021   0.0041
    13   0.0005   0.1089   0.0371   0.5009   0.1188
    14   0.0016   0.0000   0.5301   0.0133   0.3616
    15   0.0005   0.0197   0.4003   0.0257   0.4940
    16   0.0004   0.0020   0.0101   0.0046   0.0199
```

Model: Chronic 4

```
COLLINEARITY DIAGNOSTICS          VARIANCE PROPORTIONS

                     CONDITION    PORTION   PORTION   PORTION   PORTION   PORTION   PORTION   PORTION   PORTION   PORTION   PORTION   PORTION
NUMBER EIGENVALUE        INDEX   INTERCEP    O3NA01    SPNA01    S4NA01    CONA01    N2NA01  RACEWIBO   SEXMIFO   MARYINO  INCUMCON       FAT

     1     14.536        1.000     0.0000    0.0004    0.0003    0.0002    0.0007    0.0003    0.0005    0.0010    0.0011    0.0007    0.0000
     2   0.770238        4.344     0.0000    0.0007    0.0002    0.0000    0.0028    0.0006    0.0000    0.0204    0.0014    0.0000    0.0000
     3   0.544654        5.166     0.0000    0.0047    0.0019    0.0001    0.0205    0.0047    0.0004    0.5952    0.0166    0.0027    0.0000
     4   0.473015        5.539     0.0000    0.0150    0.0047    0.0009    0.0683    0.0153    0.0006    0.1056    0.1381    0.0008    0.0000
     5   0.390324        6.103     0.0000    0.0001    0.0004    0.0000    0.0008    0.0006    0.0010    0.0412    0.1486    0.0976    0.0000
     6   0.368273        6.283     0.0001    0.0057    0.0007    0.0086    0.0284    0.0044    0.0002    0.0061    0.4486    0.0002    0.0000
     7   0.196925        8.592     0.0000    0.1346    0.0075    0.0034    0.6126    0.0088    0.0107    0.0017    0.0024    0.0005    0.0000
     8   0.185356        8.856     0.0000    0.0009    0.0067    0.0086    0.0005    0.0005    0.0508    0.0083    0.1837    0.4393    0.0001
     9   0.134327       10.403     0.0000    0.0246    0.0003    0.0510    0.0000    0.0039    0.0034    0.1188    0.0411    0.0726    0.0006
    10   0.116088       11.190     0.0000    0.0727    0.0128    0.0154    0.0016    0.0138    0.6168    0.0015    0.0027    0.2067    0.0000
    11   0.080652       13.425     0.0003    0.2498    0.3284    0.0030    0.0674    0.0000    0.1813    0.0055    0.0003    0.0031    0.0000
    12   0.068077       14.613     0.0002    0.0166    0.1701    0.1446    0.0046    0.0005    0.0587    0.0022    0.0079    0.1578    0.0000
    13   0.057025       15.966     0.0000    0.3641    0.0983    0.0566    0.1416    0.5974    0.0101    0.0006    0.0044    0.0010    0.0000
    14   0.041552       18.704     0.0008    0.0421    0.0015    0.3639    0.0000    0.1470    0.0002    0.0035    0.0007    0.0143    0.0000
    15   0.023412       24.918     0.0064    0.0094    0.2057    0.1213    0.0038    0.1725    0.0295    0.0019    0.0002    0.0001    0.0003

    16   0.009679       38.754     0.0001    0.0117    0.1291    0.1901    0.0023    0.0045    0.0064    0.0005    0.0001    0.0004    0.0000
    17   0.002812       71.894     0.2410    0.0462    0.0266    0.0320    0.0414    0.0214    0.0247    0.0006    0.0000    0.0020    0.0336
    18 .000426974      184.513     0.7510    0.0006    0.0049    0.0003    0.0025    0.0037    0.0044    0.0854    0.0021    0.0002    0.9651

         PORTION   PORTION   PORTION   PORTION   PORTION   PORTION   PORTION
NUMBER     FATSQ     AGESQ  NCIGSDYN  EDCOMCON    TEMPAN    PRECAN   HUMIDAN

     1    0.0000    0.0010    0.0009    0.0002    0.0000    0.0002    0.0000
     2    0.0000    0.0144    0.8543    0.0000    0.0000    0.0000    0.0000
     3    0.0001    0.0008    0.0734    0.0001    0.0001    0.0004    0.0000
     4    0.0000    0.1317    0.0075    0.0000    0.0000    0.0137    0.0000
     5    0.0001    0.4441    0.0133    0.0046    0.0000    0.0002    0.0000
     6    0.0001    0.0547    0.0270    0.0040    0.0001    0.0314    0.0006
     7    0.0000    0.0010    0.0009    0.0000    0.0004    0.0122    0.0001
     8    0.0016    0.2163    0.0035    0.0048    0.0001    0.0067    0.0001
     9    0.0104    0.0003    0.0057    0.0001    0.0000    0.0413    0.0003
    10    0.0000    0.0423    0.0084    0.0046    0.0002    0.0083    0.0000
    11    0.0020    0.0154    0.0000    0.0508    0.0035    0.0131    0.0002
    12    0.0003    0.0077    0.0002    0.3685    0.0034    0.0005    0.0021
    13    0.0002    0.0000    0.0001    0.0064    0.0023    0.0312    0.0008
    14    0.0011    0.0400    0.0028    0.3535    0.0195    0.2276    0.0047
    15    0.0026    0.0166    0.0008    0.1675    0.0471    0.6079    0.0798
    16    0.0000    0.0003    0.0000    0.0009    0.3495    0.0015    0.5033
    17    0.0397    0.0112    0.0009    0.0296    0.5527    0.0029    0.3802
    18    0.9418    0.0023    0.0004    0.0046    0.0211    0.0007    0.0197
```

Model: Chronic 2

```
COLLINEARITY DIAGNOSTICS          VARIANCE PROPORTIONS

                       CONDITION   PORTION  PORTION  PORTION  PORTION  PORTION  PORTION  PORTION  PORTION  PORTION  PORTION  PORTION
NUMBER EIGENVALUE          INDEX  INTERCEP   D3NA01   SPNA01   S4NA01   CONA01   N2NA01 RACEWIBU  SEXMIFO MARTIDV  INCUMCON      FAI

     1     15.404          1.000    0.0000   0.0004   0.0002   0.0002   0.0006   0.0002   0.0004   0.0009   0.0009   0.0006   0.0000
     2   0.784038          4.433    0.0000   0.0003   0.0001   0.0000   0.0014   0.0003   0.0000   0.0253   0.0013   0.0001   0.0000
     3   0.553316          5.276    0.0000   0.0003   0.0000   0.0001   0.0008   0.0002   0.0001   0.4015   0.0002   0.0099   0.0000
     4   0.537131          5.355    0.0000   0.0142   0.0049   0.0000   0.0582   0.0131   0.0003   0.1805   0.0421   0.0005   0.0000
     5   0.426676          6.009    0.0000   0.0062   0.0025   0.0014   0.0326   0.0082   0.0021   0.1424   0.1653   0.0620   0.0000
     6   0.369562          6.456    0.0001   0.0056   0.0006   0.0085   0.0279   0.0040   0.0001   0.0024   0.4546   0.0029   0.0000
     7   0.196963          8.844    0.0000   0.1352   0.0081   0.0029   0.6128   0.0069   0.0091   0.0020   0.0007   0.0028   0.0000
     8   0.189620          9.013    0.0000   0.0002   0.0060   0.0094   0.0002   0.0003   0.0437   0.0079   0.1713   0.4621   0.0001
     9   0.134402         10.706    0.0000   0.0240   0.0003   0.0504   0.0000   0.0038   0.0048   0.1166   0.0352   0.0704   0.0005
    10   0.116464         11.501    0.0000   0.0720   0.0119   0.0158   0.0019   0.0139   0.6290   0.0011   0.0018   0.1919   0.0000
    11   0.081103         13.782    0.0003   0.2463   0.3164   0.0039   0.0662   0.0000   0.1802   0.0060   0.0004   0.0010   0.0000
    12   0.068322         15.016    0.0002   0.0204   0.1826   0.1421   0.0057   0.0004   0.0518   0.0022   0.0066   0.1630   0.0000
    13   0.057026         16.436    0.0000   0.3643   0.0983   0.0563   0.1416   0.5975   0.0099   0.0006   0.0039   0.0010   0.0000
    14   0.041556         19.253    0.0008   0.0422   0.0015   0.3648   0.0000   0.1473   0.0003   0.0034   0.0007   0.0143   0.0000
    15   0.023506         25.599    0.0062   0.0089   0.2070   0.1215   0.0038   0.1715   0.0280   0.0021   0.0004   0.0002   0.0003
    16   0.009682         39.887    0.0001   0.0114   0.1291   0.1890   0.0024   0.0046   0.0066   0.0005   0.0001   0.0005   0.0000
    17   0.003468         66.643    0.0603   0.0305   0.0122   0.0224   0.0165   0.0103   0.0033   0.0047   0.0520   0.0134   0.0093
    18   0.002327         81.361    0.2048   0.0172   0.0133   0.0111   0.0229   0.0120   0.0264   0.0111   0.0616   0.0032   0.0256
    19 .000425457        190.281    0.7272   0.0007   0.0049   0.0003   0.0024   0.0035   0.0036   0.0888   0.0004   0.0001   0.9640

        PORTION  PORTION  PORTION  PORTION  PORTION  PORTION  PORTION  PORTION
NUMBER    FATSQ      AGE    AGESQ NCIGSDYN EDCUMCON   TEMPAN   PRECAN  HUMIDAN

     1   0.0000   0.0000   0.0000   0.0008   0.0002   0.0000   0.0002   0.0000
     2   0.0000   0.0001   0.0009   0.7623   0.0000   0.0000   0.0000   0.0000
     3   0.0000   0.0007   0.0041   0.1496   0.0005   0.0000   0.0000   0.0000
     4   0.0001   0.0005   0.0036   0.0001   0.0011   0.0000   0.0006   0.0000
     5   0.0000   0.0003   0.0034   0.0100   0.0036   0.0001   0.0033   0.0000
     6   0.0001   0.0000   0.0002   0.0182   0.0022   0.0000   0.0113   0.0001
     7   0.0000   0.0000   0.0000   0.0009   0.0001   0.0001   0.0308   0.0006
     8   0.0016   0.0003   0.0029   0.0023   0.0055   0.0005   0.0113   0.0001
     9   0.0164   0.0000   0.0000   0.0053   0.0001   0.0001   0.0082   0.0001
    10   0.0000   0.0001   0.0006   0.0073   0.0039   0.0000   0.0410   0.0003
    11   0.0020   0.0002   0.0000   0.0001   0.0564   0.0002   0.0084   0.0000
    12   0.0002   0.0001   0.0000   0.0001   0.3537   0.0035   0.0135   0.0002
    13   0.0002   0.0000   0.0000   0.0001   0.0065   0.0031   0.0006   0.0020
    14   0.0011   0.0000   0.0010   0.0026   0.3465   0.0023   0.0312   0.0008
    15   0.0026   0.0006   0.0023   0.0011   0.1707   0.0196   0.2283   0.0047
    16   0.0000   0.0002   0.0003   0.0000   0.0008   0.0455   0.6066   0.0776
    17   0.0166   0.4720   0.4848   0.0241   0.0414   0.3457   0.0015   0.5067
    18   0.0222   0.5207   0.4929   0.0143   0.0012   0.3651   0.0009   0.1909
    19   0.9428   0.0042   0.0031   0.0010   0.0055   0.2534   0.0016   0.1967
                                                       0.0208   0.0007   0.0191
```

Model: Acute 26

```
COLLINEARITY DIAGNOSTICS        VARIANCE PROPORTIONS

                     CONDITION    PORTION   PORTION  PORTION   PORTION   PORTION   PORTION   PORTION   PORTION   PORTION   PORTION   PORTION
NUMBER EIGENVALUE        INDEX   INTERCEP    03HR01 03AVMAXT    SPHR01    S4HR01  RACEWIBO   SEXMIFO   MARYIND  INCOMCON  CITYI100       FAT

     1     15.437        1.000     0.0000    0.0000   0.0000    0.0005    0.0006    0.0004    0.0009    0.0008    0.0005    0.0009    0.0000
     2   0.994133        3.941     0.0000    0.0000   0.0000    0.0000    0.0002    0.0002    0.0038    0.0006    0.0047    0.0009    0.0000
     3   0.761458        4.503     0.0000    0.0004   0.0005    0.0010    0.0004    0.0000    0.0305    0.0003    0.0004    0.0000    0.0000
     4   0.654337        4.857     0.0000    0.0020   0.0028    0.0044    0.0060    0.0010    0.0155    0.0171    0.0050    0.0001    0.0000
     5   0.565559        5.224     0.0000    0.0002   0.0002    0.0015    0.0027    0.0018    0.1236    0.0325    0.0069    0.5013    0.0000
     6   0.516000        5.470     0.0000    0.0000   0.0000    0.0001    0.0014    0.0029    0.5050    0.0183    0.0009    0.0961    0.0000
     7   0.440225        5.922     0.0000    0.0001   0.0001    0.0008    0.0051    0.0000    0.0088    0.0526    0.0017    0.0497    0.0000
     8   0.373058        6.433     0.0000    0.0003   0.0006    0.0026    0.0032    0.0085    0.0595    0.2051    0.0248    0.0008    0.0000
     9   0.302360        7.145     0.0001    0.0007   0.0014    0.0244    0.0266    0.0143    0.0072    0.3679    0.0017    0.1372    0.0000
    10   0.212851        8.516     0.0000    0.0000   0.0000    0.0178    0.0676    0.0333    0.0086    0.0463    0.2331    0.0896    0.0000
    11   0.191200        8.985     0.0000    0.0001   0.0001    0.0357    0.5027    0.0087    0.0001    0.0464    0.0042    0.0033    0.0000
    12   0.142263       10.417     0.0000    0.0000   0.0000    0.0026    0.0073    0.0092    0.0938    0.0124    0.1053    0.0134    0.0006
    13   0.135885       10.658     0.0000    0.0002   0.0003    0.5759    0.1499    0.0677    0.0136    0.0155    0.0061    0.0011    0.0000
    14   0.114699       11.601     0.0001    0.0001   0.0000    0.2190    0.0439    0.4330    0.0048    0.0267    0.3374    0.0141    0.0000
    15   0.073556       14.487     0.0010    0.0000   0.0002    0.0242    0.0211    0.3384    0.0094    0.0122    0.1655    0.0439    0.0000
    16   0.043085       18.929     0.0000    0.0234   0.0006    0.0157    0.0174    0.0001    0.0005    0.0018    0.0002    0.0158    0.0000
    17   0.030306       22.569     0.0049    0.0008   0.0013    0.0027    0.0381    0.0441    0.0017    0.0038    0.0067    0.0082    0.0001
    18   0.006478       48.814     0.0978    0.0546   0.0383    0.0621    0.0074    0.0271    0.0008    0.0026    0.0055    0.0204    0.0077
    19   0.002873       73.300     0.0000    0.5014   0.5118    0.0055    0.0074    0.0058    0.0078    0.0637    0.0052    0.0009    0.0081
    20   0.002267       82.523     0.1241    0.4140   0.4394    0.0012    0.0166    0.0078    0.0090    0.0743    0.0041    0.0021    0.0103
    21 .000426347      190.283     0.7720    0.0018   0.0025    0.0023    0.0004    0.0036    0.0863    0.0008    0.0001    0.0003    0.9810

           PORTION   PORTION   PORTION   PORTION   PORTION   PORTION   PORTION   PORTION   PORTION   PORTION
NUMBER       FATSQ       AGE     AGESQ  NCIGSDYN  EDCOMCON  CHRLMDUM  CROWDING  AVMAXTMP  AVPRECIP   HUMIDRF

     1      0.0000    0.0000    0.0000    0.0008    0.0002    0.0005    0.0006    0.0001    0.0009    0.0001
     2      0.0000    0.0002    0.0014    0.0329    0.0004    0.4809    0.0035    0.0000    0.0000    0.0000
     3      0.0000    0.0000    0.0000    0.6971    0.0000    0.0438    0.0002    0.0002    0.0027    0.0000
     4      0.0000    0.0001    0.0002    0.1009    0.0005    0.0368    0.0000    0.0003    0.1046    0.0002
     5      0.0000    0.0000    0.0001    0.0027    0.0000    0.0088    0.0073    0.0000    0.0045    0.0001
     6      0.0001    0.0002    0.0009    0.0866    0.0001    0.0211    0.0021    0.0000    0.0336    0.0000
     7      0.0000    0.0008    0.0051    0.0147    0.0007    0.2097    0.0233    0.0001    0.2529    0.0001
     8      0.0000    0.0003    0.0030    0.0037    0.0008    0.1430    0.0561    0.0001    0.2078    0.0000
     9      0.0001    0.0001    0.0001    0.0021    0.0034    0.0053    0.0184    0.0000    0.1954    0.0011
    10      0.0001    0.0000    0.0006    0.0001    0.0181    0.0337    0.2710    0.0003    0.0058    0.0000
    11      0.0004    0.0002    0.0009    0.0013    0.0001    0.0099    0.1820    0.0007    0.0028    0.0001
    12      0.0097    0.0003    0.0035    0.0000    0.0006    0.0004    0.1906    0.0000    0.0002    0.0001
    13      0.0011    0.0001    0.0007    0.0051    0.0010    0.0000    0.0359    0.0003    0.0674    0.0001
    14      0.0004    0.0001    0.0012    0.0104    0.0041    0.0013    0.0024    0.0000    0.0359    0.0020
    15      0.0026    0.0003    0.0001    0.0000    0.2937    0.0003    0.0075    0.0066    0.0311    0.0137
    16      0.0000    0.0000    0.0002    0.0003    0.0304    0.0003    0.0054    0.2567    0.0167    0.0248
    17      0.0027    0.0000    0.0053    0.0039    0.5334    0.0020    0.1513    0.0068    0.0279    0.2374
    18      0.0138    0.0340    0.0484    0.0051    0.1048    0.0003    0.0297    0.0166    0.0093    0.6951
    19      0.0009    0.4480    0.4404    0.0204    0.0027    0.0003    0.0055    0.3269    0.0000    0.0099
    20      0.0073    0.5105    0.4847    0.0118    0.0000    0.0005    0.0146    0.3790    0.0012    0.0077
    21      0.9609    0.0049    0.0031    0.0002    0.0050    0.0012    0.0008    0.0054    0.0000    0.0075
```

```
COLLINEARITY DIAGNOSTICS        VARIANCE PROPORTIONS

                     CONDITION   PORTION   PORTION   PORTION   PORTION   PORTION   PORTION   PORTION   PORTION   PORTION   PORTION   PORTION
NUMBER EIGENVALUE        INDEX  INTERCEP    O3NR01    SPNR01    S4NR01    CONR01    N2NR01  RACEMIBO   SEXMIFO   MARYING  INCOMCON  CITYIJUO

     1     17.192        1.000    0.0000    0.0004    0.0004    0.0005    0.0007    0.0004    0.0003    0.0006    0.0006    0.0004    0.0008
     2   0.946923        4.261    0.0000    0.0007    0.0001    0.0004    0.0002    0.0001    0.0001    0.0020    0.0001    0.0005    0.0000
     3   0.656742        5.116    0.0000    0.0000    0.0012    0.0001    0.0013    0.0006    0.0012    0.0080    0.0497    0.0045    0.1813
     4   0.575899        5.464    0.0000    0.0071    0.0047    0.0007    0.0131    0.0054    0.0006    0.0708    0.0257    0.0017    0.2044
     5   0.518186        5.760    0.0000    0.0492    0.0103    0.0000    0.0041    0.0073    0.0005    0.0722    0.0034    0.0001    0.0999
     6   0.484490        5.957    0.0000    0.0038    0.0011    0.0008    0.0056    0.0004    0.0001    0.1096    0.0422    0.0000    0.0316
     7   0.478519        5.994    0.0000    0.0067    0.0005    0.0079    0.0097    0.0035    0.0034    0.2194    0.0699    0.0010    0.2521
     8   0.355070        6.958    0.0000    0.0649    0.0018    0.0143    0.5672    0.0073    0.0002    0.0305    0.0179    0.0000    0.0686
     9   0.312230        7.420    0.0000    0.0188    0.0003    0.0097    0.0187    0.0002    0.0000    0.1056    0.3816    0.0049    0.0097
    10   0.304593        7.513    0.0000    0.0243    0.0050    0.0008    0.0449    0.0045    0.0002    0.0187    0.0282    0.0173    0.0009
    11   0.212549        8.993    0.0001    0.1665    0.0022    0.0031    0.0389    0.0138    0.0319    0.0281    0.1368    0.0019    0.0263
    12   0.184467        9.654    0.0000    0.0015    0.0066    0.0001    0.0088    0.0006    0.1488    0.0197    0.0057    0.1396    0.0907
    13   0.168892       10.089    0.0000    0.0441    0.0853    0.5834    0.0194    0.0560    0.0199    0.0008    0.0392    0.0066    0.0028
    14   0.127682       11.604    0.0000    0.0086    0.0659    0.0966    0.0262    0.0126    0.0000    0.0937    0.0238    0.1814    0.0168
    15   0.122933       11.826    0.0000    0.0173    0.2456    0.1429    0.0536    0.4981    0.0620    0.0063    0.0044    0.0043    0.0020
    16   0.121365       11.902    0.0000    0.0825    0.0991    0.0002    0.1087    0.1978    0.1734    0.0597    0.0242    0.1579    0.0000
    17   0.107294       12.658    0.0000    0.0410    0.3997    0.0713    0.0675    0.0765    0.1946    0.0006    0.0115    0.3419    0.0043
    18   0.059007       17.069    0.0017    0.0693    0.0163    0.0391    0.0093    0.0147    0.2778    0.0142    0.0093    0.1074    0.0319
    19   0.038059       21.254    0.0001    0.3555    0.0014    0.0666    0.0005    0.0573    0.0097    0.0006    0.0001    0.0008    0.0163
    20   0.025252       26.092    0.0043    0.0020    0.0051    0.0219    0.0000    0.0348    0.0246    0.0849    0.0007    0.0144    0.0038
    21   0.005964       53.690    0.1209    0.0321    0.0455    0.0002    0.0006    0.0066    0.0327    0.0008    0.0004    0.0059    0.0152
    22   0.001880       95.623    0.0939    0.0005    0.0003    0.0001    0.0002    0.0000    0.0100    0.0096    0.1242    0.0001    0.0004
    23 .000421528      201.950    0.7798    0.0010    0.0024    0.0001    0.0008    0.0014    0.0057    0.1241    0.0004    0.0004    0.0002

         PORTION   PORTION   PORTION   PORTION   PORTION   PORTION   PORTION   PORTION   PORTION   PORTION   PORTION   PORTION
NUMBER       FAT     FATSQ       AGE     AGESQ  SMOKYING  EDCOMCON  CHRLHOUN  CROWDING  AVMAXTMP  AVPRECIP   HUMIDRF  INDUCCAV

     1    0.0000    0.0000    0.0000    0.0000    0.0008    0.0001    0.0003    0.0005    0.0001    0.0007    0.0000    0.0006
     2    0.0000    0.0000    0.0000    0.0003    0.0079    0.0000    0.8380    0.0009    0.0000    0.0001    0.0000    0.0001
     3    0.0000    0.0000    0.0001    0.0005    0.4220    0.0000    0.0199    0.0010    0.0000    0.0054    0.0000    0.0015
     4    0.0000    0.0000    0.0000    0.0000    0.3670    0.0000    0.0126    0.0001    0.0004    0.0440    0.0000    0.0070
     5    0.0000    0.0000    0.0002    0.0014    0.0018    0.0000    0.0348    0.0065    0.0007    0.2425    0.0000    0.0097
     6    0.0000    0.0000    0.0004    0.0037    0.1249    0.0002    0.0266    0.0182    0.0000    0.1645    0.0001    0.0002
     7    0.0000    0.0001    0.0001    0.0006    0.0034    0.0012    0.0044    0.0808    0.0006    0.0700    0.0002    0.0115
     8    0.0000    0.0000    0.0000    0.0009    0.0001    0.0001    0.0005    0.0121    0.0028    0.0005    0.0000    0.0000
     9    0.0000    0.0000    0.0003    0.0037    0.0001    0.0001    0.0241    0.1068    0.0004    0.0135    0.0000    0.0089
    10    0.0000    0.0000    0.0000    0.0003    0.0076    0.0049    0.0000    0.0157    0.0004    0.1271    0.0000    0.4574
    11    0.0001    0.0007    0.0002    0.0008    0.0062    0.0000    0.0096    0.1932    0.0009    0.1305    0.0010    0.0832
    12    0.0000    0.0001    0.0001    0.0020    0.0032    0.0074    0.0153    0.2467    0.0014    0.0155    0.0003    0.1656
    13    0.0000    0.0000    0.0001    0.0007    0.0015    0.0000    0.0042    0.0570    0.0011    0.0120    0.0001    0.0008
    14    0.0005    0.0073    0.0001    0.0025    0.0003    0.0007    0.0050    0.1629    0.0001    0.0079    0.0005    0.0012
    15    0.0000    0.0005    0.0000    0.0001    0.0004    0.0000    0.0003    0.0024    0.0058    0.0301    0.0013    0.0176
    16    0.0002    0.0042    0.0000    0.0001    0.0027    0.0002    0.0004    0.0000    0.0061    0.0086    0.0002    0.0111
    17    0.0000    0.0001    0.0000    0.0001    0.0000    0.0001    0.0000    0.0028    0.0030    0.0279    0.0008    0.0180
    18    0.0000    0.0041    0.0005    0.0000    0.0059    0.1812    0.0027    0.0061    0.1220    0.0393    0.0176    0.1286
    19    0.0000    0.0000    0.0001    0.0000    0.0001    0.1752    0.0000    0.0041    0.6494    0.0022    0.0295    0.0462
    20    0.0001    0.0030    0.0003    0.0033    0.0096    0.5061    0.0005    0.1137    0.0021    0.0248    0.2945    0.0142
    21    0.0094    0.0176    0.0375    0.0489    0.0155    0.1127    0.0001    0.0355    0.1860    0.0042    0.6329    0.0039
    22    0.0089    0.0062    0.9600    0.9304    0.0193    0.0002    0.0001    0.0113    0.0108    0.0000    0.0131    0.0117
    23    0.9805    0.9560    0.0000    0.0004    0.0001    0.0076    0.0008    0.0013    0.0056    0.0000    0.0058    0.0012
```

Model: Acute 8

```
COLLINEARITY DIAGNOSTICS         VARIANCE PROPORTIONS

                        CONDITION   PORTION  PORTION  PORTION  PORTION  PORTION  PORTION  PORTION  PORTION  PORTION  PORTION  PORTION
NUMBER EIGENVALUE           INDEX  INTERCEP O3NR01SQ   SPNR01   S4NR01   CONR01   N2NR01 WACEW100  SEXMIFO  MARYINO INCOMCON CITYI100

     1     15.864           1.000    0.0000   0.0006   0.0005   0.0006   0.0008   0.0005   0.0004   0.0008   0.0008   0.0005   0.0009
     2   0.987611           4.008    0.0000   0.0042   0.0001   0.0002   0.0002   0.0003   0.0002   0.0035   0.0008   0.0049   0.0008
     3   0.810936           4.423    0.0000   0.5337   0.0028   0.0000   0.0004   0.0024   0.0001   0.0010   0.0004   0.0004   0.0097
     4   0.669048           4.869    0.0000   0.0028   0.0007   0.0000   0.0003   0.0001   0.0021   0.0005   0.0260   0.0067   0.0753
     5   0.590181           5.185    0.0000   0.0007   0.0035   0.0024   0.0144   0.0031   0.0007   0.2506   0.0237   0.0057   0.2549
     6   0.516269           5.543    0.0000   0.0020   0.0001   0.0013   0.0001   0.0001   0.0086   0.4220   0.0186   0.0015   0.1609
     7   0.435901           6.033    0.0000   0.0206   0.0004   0.0083   0.0090   0.0000   0.0001   0.0088   0.0460   0.0012   0.0233
     8   0.400807           6.291    0.0000   0.0777   0.0047   0.0064   0.3203   0.0132   0.0020   0.0065   0.0034   0.0069   0.1074
     9   0.362433           6.616    0.0000   0.0064   0.0000   0.0009   0.0900   0.0050   0.0002   0.0497   0.3572   0.0183   0.0467
    10   0.285156           7.459    0.0001   0.0933   0.0108   0.0347   0.2692   0.0014   0.0117   0.0071   0.2057   0.0016   0.0027
    11   0.217720           8.536    0.0000   0.0020   0.0145   0.0026   0.0133   0.0001   0.0225   0.0122   0.0300   0.2540   0.1176
    12   0.180233           9.382    0.0000   0.0284   0.1158   0.5171   0.0090   0.0091   0.0150   0.0120   0.0578   0.0120   0.0037
    13   0.142593          10.549    0.0000   0.0038   0.0000   0.0194   0.0037   0.0165   0.0135   0.0961   0.0041   0.0708   0.0129
    14   0.137285          10.750    0.0000   0.0636   0.1438   0.1152   0.1931   0.2813   0.0720   0.0015   0.0325   0.1622   0.0022
    15   0.122212          11.393    0.0000   0.0147   0.2864   0.1195   0.0296   0.5130   0.1170   0.0011   0.0044   0.0005   0.0106
    16   0.112021          11.900    0.0000   0.0216   0.3024   0.0991   0.0326   0.0229   0.4072   0.0022   0.0187   0.2681   0.0182
    17   0.074604          14.582    0.0007   0.0015   0.0325   0.0352   0.0003   0.0628   0.2333   0.0126   0.0085   0.1596   0.0298
    18   0.053908          17.155    0.0001   0.0803   0.0346   0.0008   0.0063   0.0157   0.0072   0.0008   0.0006   0.0101   0.0038
    19   0.028041          23.785    0.0067   0.0009   0.0001   0.0354   0.0001   0.0448   0.0529   0.0036   0.0016   0.0023   0.0056
    20   0.005987          51.477    0.1240   0.0330   0.0438   0.0001   0.0053   0.0042   0.0254   0.0018   0.0030   0.0080   0.0230
    21   0.002542          78.996    0.0724   0.0028   0.0004   0.0005   0.0006   0.0000   0.0097   0.0159   0.1336   0.0078   0.0001
    22 .000414016         195.749    0.7940   0.0005   0.0020   0.0007   0.0014   0.0026   0.0034   0.0892   0.0005   0.0003   0.0000

        PORTION  PORTION  PORTION  PORTION  PORTION  PORTION  PORTION  PORTION  PORTION  PORTION  PORTION
NUMBER      FAT    FATSQ      AGE    AGESQ SMOKYING EDCOMCON CHALMOUN CROWDING AVMAXTMP AVPRECIP  HUMIDAF

     1   0.0000   0.0000   0.0000   0.0000   0.0009   0.0002   0.0005   0.0006   0.0002   0.0009   0.0001
     2   0.0000   0.0000   0.0002   0.0015   0.0200   0.0005   0.4920   0.0039   0.0001   0.0001   0.0000
     3   0.0000   0.0000   0.0000   0.0000   0.0136   0.0001   0.0079   0.0003   0.0003   0.0507   0.0000
     4   0.0000   0.0000   0.0001   0.0006   0.5666   0.0002   0.0648   0.0022   0.0000   0.0096   0.0000
     5   0.0000   0.0000   0.0000   0.0000   0.1101   0.0000   0.0512   0.0033   0.0003   0.0009   0.0000
     6   0.0000   0.0001   0.0001   0.0005   0.1523   0.0001   0.0032   0.0022   0.0008   0.0268   0.0000
     7   0.0000   0.0000   0.0006   0.0042   0.0574   0.0007   0.1607   0.0326   0.0004   0.2759   0.0002
     8   0.0000   0.0000   0.0002   0.0016   0.0131   0.0012   0.0663   0.0071   0.0001   0.1436   0.0000
     9   0.0000   0.0000   0.0002   0.0021   0.0095   0.0000   0.0952   0.0547   0.0000   0.0682   0.0000
    10   0.0000   0.0002   0.0001   0.0002   0.0000   0.0017   0.0059   0.0269   0.0025   0.2533   0.0004
    11   0.0000   0.0001   0.0001   0.0010   0.0000   0.0174   0.0420   0.3271   0.0003   0.0069   0.0000
    12   0.0000   0.0002   0.0002   0.0012   0.0035   0.0000   0.0045   0.1354   0.0001   0.0012   0.0003
    13   0.0006   0.0103   0.0002   0.0029   0.0001   0.0004   0.0011   0.1283   0.0001   0.0001   0.0001
    14   0.0000   0.0004   0.0002   0.0015   0.0026   0.0029   0.0004   0.0763   0.0006   0.0297   0.0002
    15   0.0000   0.0002   0.0000   0.0000   0.0013   0.0001   0.0000   0.0014   0.0042   0.0252   0.0008
    16   0.0000   0.0001   0.0001   0.0010   0.0029   0.0017   0.0015   0.0002   0.0002   0.0377   0.0012
    17   0.0000   0.0024   0.0002   0.0000   0.0021   0.2516   0.0000   0.0053   0.1368   0.0224   0.0042
    18   0.0000   0.0000   0.0002   0.0003   0.0001   0.2153   0.0000   0.0228   0.6478   0.0003   0.0125
    19   0.0002   0.0036   0.0002   0.0055   0.0120   0.3999   0.0007   0.1222   0.0039   0.0398   0.2081
    20   0.0096   0.0166   0.0479   0.0645   0.0120   0.0998   0.0000   0.0250   0.1947   0.0068   0.6615
    21   0.0049   0.0021   0.9460   0.9095   0.0191   0.0001   0.0012   0.0206   0.0134   0.0001   0.0230
    22   0.9848   0.9639   0.0032   0.0020   0.0007   0.0062   0.0007   0.0016   0.0044   0.0000   0.0073
```

```
COLLINEARITY DIAGNOSTICS          VARIANCE PROPORTIONS

                            CONDITION   PORTION   PORTION   PORTION   PORTION   PORTION   PORTION   PORTION   PORTION   PORTION   PORTION   PORTION
NUMBER EIGENVALUE               INDEX  INTERCEP    O3NR01    SPNR01    S4NR01    CONR01    N2NR01  RACEWIBO   SEXMIFO  MARYIND  INCOMCON  CITYI100

     1     16.313               1.000    0.0000    0.0006    0.0004    0.0005    0.0008    0.0004    0.0006    0.0008    0.0007    0.0005    0.0008
     2   0.986203               4.067    0.0000    0.0005    0.0000    0.0002    0.0002    0.0001    0.0002    0.0035    0.0008    0.0051    0.0005
     3   0.670512               4.933    0.0000    0.0006    0.0003    0.0000    0.0003    0.0000    0.0020    0.0001    0.0235    0.0061    0.0780
     4   0.593550               5.243    0.0000    0.0033    0.0054    0.0022    0.0132    0.0044    0.0009    0.2442    0.0270    0.0065    0.2332
     5   0.510939               5.607    0.0000    0.0041    0.0017    0.0017    0.0001    0.0013    0.0020    0.4003    0.0152    0.0011    0.0921
     6   0.489473               5.773    0.0000    0.0503    0.0113    0.0005    0.0109    0.0075    0.0012    0.0315    0.0070    0.0000    0.1834
     7   0.431102               6.151    0.0000    0.0094    0.0007    0.0048    0.0013    0.0021    0.0001    0.0187    0.0632    0.0038    0.0965
     8   0.372406               6.619    0.0000    0.0315    0.0000    0.0072    0.5015    0.0105    0.0007    0.0176    0.1053    0.0000    0.0408
     9   0.349332               6.834    0.0000    0.0489    0.0021    0.0240    0.0973    0.0803    0.0002    0.0227    0.3868    0.0177    0.0479
    10   0.227902               8.461    0.0001    0.1658    0.0056    0.0000    0.0692    0.0148    0.0378    0.0305    0.1697    0.0865    0.0016
    11   0.216634               8.678    0.0000    0.0252    0.0076    0.0009    0.0025    0.0012    0.0085    0.0031    0.0025    0.1827    0.1298
    12   0.179186               9.650    0.0000    0.0566    0.0914    0.5960    0.0139    0.0420    0.0099    0.0035    0.0233    0.0026    0.0061
    13   0.142313              10.707    0.0000    0.0011    0.0034    0.0446    0.0002    0.0007    0.0068    0.0911    0.0089    0.1068    0.0144
    14   0.130356              11.187    0.0000    0.0948    0.2138    0.0384    0.1660    0.1448    0.1941    0.0062    0.0205    0.1394    0.0000
    15   0.122145              11.557    0.0000    0.0230    0.2230    0.1228    0.0566    0.5029    0.0613    0.0007    0.0089    0.0006    0.0082
    16   0.111409              12.101    0.0000    0.0311    0.3728    0.0837    0.0577    0.0462    0.3281    0.0024    0.0176    0.2568    0.0155
    17   0.072736              14.976    0.0009    0.0135    0.0059    0.0320    0.0030    0.0398    0.2558    0.0131    0.0080    0.1639    0.0327
    18   0.039830              20.238    0.0002    0.3814    0.0096    0.0003    0.0011    0.0344    0.0019    0.0011    0.0034    0.0007    0.0237
    19   0.027740              24.250    0.0062    0.0227    0.0001    0.0385    0.0000    0.0618    0.0494    0.0030    0.0009    0.0023    0.0023
    20   0.006000              52.141    0.1236    0.0329    0.0423    0.0000    0.0046    0.0033    0.0261    0.0010    0.0033    0.0074    0.0223
    21   0.002543              80.090    0.0720    0.0028    0.0003    0.0005    0.0004    0.0000    0.0098    0.0189    0.1330    0.0077    0.0001
    22 .000414017             198.500    0.7968    0.0005    0.0020    0.0007    0.0014    0.0024    0.0034    0.0891    0.0005    0.0003    0.0000

          PORTION   PORTION   PORTION   PORTION   PORTION   PORTION   PORTION   PORTION   PORTION   PORTION   PORTION
NUMBER        FAT     FATSQ       AGE     AGESQ  SMOKYIND  EDCOMCON  CHALHOUR  CROWDING  AVMAXTMP  AVPRECIP   HUMIDRF

     1     0.0000    0.0000    0.0000    0.0000    0.0009    0.0002    0.0005    0.0005    0.0001    0.0000    0.0001
     2     0.0000    0.0000    0.0002    0.0015    0.0211    0.0005    0.5007    0.0040    0.0000    0.0006    0.0000
     3     0.0000    0.0000    0.0001    0.0006    0.5909    0.0002    0.0617    0.0022    0.0000    0.0030    0.0000
     4     0.0000    0.0000    0.0000    0.0000    0.1034    0.0000    0.0399    0.0034    0.0004    0.0110    0.0000
     5     0.0000    0.0001    0.0001    0.0006    0.1232    0.0001    0.0079    0.0018    0.0000    0.1090    0.0001
     6     0.0000    0.0000    0.0000    0.0000    0.0297    0.0000    0.0034    0.0004    0.0006    0.3638    0.0001
     7     0.0000    0.0000    0.0009    0.0058    0.0620    0.0012    0.2358    0.0432    0.0006    0.0637    0.0001
     8     0.0000    0.0000    0.0000    0.0002    0.0001    0.0002    0.0115    0.0210    0.0011    0.0048    0.0000
     9     0.0000    0.0000    0.0002    0.0020    0.0108    0.0001    0.0852    0.0210    0.0016    0.0185    0.0000
    10     0.0001    0.0005    0.0001    0.0000    0.0001    0.0089    0.0002    0.0201    0.0005    0.2474    0.0011
    11     0.0000    0.0003    0.0002    0.0013    0.0000    0.0117    0.0452    0.4336    0.0008    0.0036    0.0000
    12     0.0000    0.0000    0.0001    0.0008    0.0036    0.0000    0.0083    0.0609    0.0020    0.0242    0.0000
    13     0.0006    0.0093    0.0002    0.0035    0.0000    0.0007    0.0009    0.1618    0.0003    0.0009    0.0001
    14     0.0000    0.0010    0.0001    0.0010    0.0056    0.0040    0.0007    0.0270    0.0028    0.0106    0.0006
    15     0.0000    0.0001    0.0000    0.0000    0.0005    0.0003    0.0001    0.0018    0.0088    0.0243    0.0013
    16     0.0000    0.0001    0.0001    0.0009    0.0021    0.0014    0.0015    0.0005    0.0019    0.0370    0.0006
    17     0.0000    0.0026    0.0003    0.0000    0.0027    0.3433    0.0000    0.0022    0.0319    0.0289    0.0081
    18     0.0000    0.0003    0.0001    0.0009    0.0010    0.1692    0.0003    0.0373    0.7137    0.0089    0.0022
    19     0.0002    0.0033    0.0002    0.0049    0.0108    0.3507    0.0006    0.1084    0.0303    0.0326    0.3047
    20     0.0093    0.0165    0.0474    0.0640    0.0116    0.1009    0.0000    0.0262    0.1858    0.0068    0.6516
    21     0.0049    0.0020    0.9465    0.9100    0.0193    0.0001    0.0011    0.0209    0.0131    0.0001    0.0222
    22     0.9849    0.9640    0.0032    0.0020    0.0007    0.0062    0.0006    0.0017    0.0041    0.0000    0.0072
```

Model: Acute 3

```
COLLINEARITY DIAGNOSTICS          VARIANCE PROPORTIONS

                      CONDITION   PORTION  PORTION  PORTION  PORTION  PORTION  PORTION  PORTION  PORTION  PORTION  PORTION  PORTION
NUMBER EIGENVALUE         INDEX  INTERCEP   O3HR01   SPHR01   SAHR01 RACEW100 SEXMIFO  MARYING INCOMCON CITYI100      FAT    FATSQ

     1     14.867         1.000    0.0000   0.0006   0.0005   0.0007   0.0004   0.0010   0.0009   0.0006   0.0010   0.0000   0.0000
     2   0.991263         3.873    0.0000   0.0004   0.0000   0.0002   0.0002   0.0036   0.0009   0.0050   0.0003   0.0000   0.0000
     3   0.657897         4.754    0.0000   0.0007   0.0001   0.0000   0.0016   0.0054   0.0204   0.0051   0.0445   0.0000   0.0000
     4   0.568695         5.113    0.0000   0.0023   0.0081   0.0042   0.0007   0.3328   0.0184   0.0082   0.2950   0.0000   0.0000
     5   0.505217         5.425    0.0000   0.0005   0.0012   0.0011   0.0029   0.3390   0.0167   0.0020   0.2187   0.0000   0.0001
     6   0.466162         5.647    0.0000   0.0508   0.0173   0.0000   0.0008   0.0159   0.0113   0.0004   0.0679   0.0000   0.0000
     7   0.421673         5.938    0.0000   0.0176   0.0026   0.0057   0.0001   0.0185   0.0329   0.0057   0.0847   0.0000   0.0000
     8   0.349887         6.519    0.0000   0.0295   0.0083   0.0208   0.0010   0.0220   0.4898   0.0158   0.0894   0.0000   0.0000
     9   0.231682         8.011    0.0001   0.2786   0.0015   0.0056   0.0328   0.0282   0.1588   0.0617   0.0083   0.0001   0.0004
    10   0.209470         8.425    0.0000   0.0361   0.0086   0.0318   0.0187   0.0021   0.0100   0.1849   0.1070   0.0000   0.0003
    11   0.186737         8.923    0.0000   0.0566   0.0307   0.6748   0.0073   0.0048   0.0273   0.0038   0.0000   0.0000   0.0001
    12   0.141863        10.237    0.0000   0.0001   0.0043   0.0084   0.0097   0.0940   0.0134   0.1014   0.0126   0.0006   0.0098
    13   0.133133        10.568    0.0000   0.0648   0.6072   0.1169   0.0823   0.0118   0.0130   0.0755   0.0013   0.0000   0.0010
    14   0.114856        11.377    0.0001   0.0039   0.2223   0.0443   0.4324   0.0038   0.0277   0.3411   0.0147   0.0000   0.0003
    15   0.072640        14.306    0.0010   0.0177   0.0192   0.0199   0.3187   0.0106   0.0122   0.1724   0.0442   0.0000   0.0026
    16   0.042719        18.655    0.0000   0.3964   0.0204   0.0169   0.0000   0.0006   0.0013   0.0000   0.0154   0.0000   0.0000
    17   0.029774        22.346    0.0053   0.0018   0.0005   0.0443   0.0463   0.0024   0.0031   0.0055   0.0070   0.0001   0.0027
    18   0.006072        49.482    0.1259   0.0385   0.0492   0.0029   0.0278   0.0011   0.0042   0.0062   0.0214   0.0095   0.0168
    19   0.002532        76.632    0.0767   0.0024   0.0013   0.0011   0.0124   0.0140   0.1371   0.0076   0.0004   0.0048   0.0020
    20 .000422746       187.532    0.7908   0.0010   0.0028   0.0002   0.0042   0.0886   0.0008   0.0001   0.0001   0.9849   0.9641

        PORTION  PORTION  PORTION  PORTION  PORTION  PORTION  PORTION  PORTION  PORTION
NUMBER      AGE    AGESQ SMOKYING EDCOMCON CHRLHOUM CROWDING AVMAXTMP AVPRECIP  HUMIDRF

     1   0.0000   0.0000   0.0010   0.0002   0.0006   0.0006   0.0002   0.0010   0.0001
     2   0.0002   0.0016   0.0228   0.0005   0.4926   0.0042   0.0000   0.0003   0.0000
     3   0.0001   0.0005   0.6484   0.0002   0.0759   0.0010   0.0000   0.0029   0.0000
     4   0.0000   0.0000   0.0187   0.0000   0.0311   0.0042   0.0005   0.0021   0.0000
     5   0.0001   0.0003   0.1532   0.0001   0.0004   0.0016   0.0000   0.0798   0.0001
     6   0.0000   0.0001   0.0136   0.0000   0.0001   0.0002   0.0008   0.5223   0.0001
     7   0.0011   0.0071   0.0782   0.0011   0.2752   0.0504   0.0010   0.0157   0.0001
     8   0.0001   0.0014   0.0002   0.0000   0.0748   0.0190   0.0012   0.0129   0.0001
     9   0.0001   0.0000   0.0008   0.0090   0.0008   0.0136   0.0036   0.1725   0.0016
    10   0.0001   0.0009   0.0011   0.0133   0.0359   0.3789   0.0019   0.0002   0.0000
    11   0.0001   0.0006   0.0022   0.0001   0.0066   0.0960   0.0052   0.0186   0.0000
    12   0.0003   0.0036   0.0000   0.0005   0.0003   0.1981   0.0000   0.0002   0.0001
    13   0.0001   0.0007   0.0061   0.0012   0.0000   0.0231   0.0000   0.0505   0.0001
    14   0.0001   0.0013   0.0049   0.0035   0.0013   0.0025   0.0080   0.0341   0.0020
    15   0.0003   0.0001   0.0041   0.3144   0.0001   0.0055   0.0148   0.0321   0.0138
    16   0.0000   0.0002   0.0002   0.0461   0.0004   0.0084   0.7218   0.0168   0.0232
    17   0.0001   0.0053   0.0130   0.4939   0.0019   0.1418   0.0032   0.0284   0.2731
    18   0.0402   0.0656   0.0123   0.1103   0.0002   0.0287   0.2286   0.0091   0.6516
    19   0.9441   0.9074   0.0184   0.0005   0.0008   0.0211   0.0140   0.0006   0.0256
    20   0.0050   0.0033   0.0008   0.0049   0.0012   0.0008   0.0030   0.0000   0.0085
```

Footnotes to Chapter 4

[1] See, for instance, d'Arge, Kask, Case, Ben-David, Eubanks, and Anderson [1982].

[2] For a good discussion of possible non-linearities in air pollution dose-response estimation, see Graves and Krumm [1981] and Krumm and Graves [1983].

[3] For an interesting application of logit analysis to asthmatic morbidity, see Whittemore and Korn [1980].

[4] Some epidemiologic investigations of ozone have attempted to control for pollen concentrations. See Zagraniski et al. [1979].

[5] See Graves and Krumm [1981] for a discussion of interactive effects. See also Bates et al. [1973] for an analysis of ozone interactive effects with other pollutants.

[6] This is the purpose of many of the studies cited in Environmental Protection Agency [1978] and National Academy of Sciences [1977].

[7] See Lave and Seskin [1977], pp. 44-50.

[7a] For example, see McDonnell et al. [1983].

[8] Using an earlier version of the Health Interview Survey, Ostro [1984] also found such a relationship between particulate matter and restricted activity days.

[9] See Friedman [1981], for instance.

[10] Others investigating ozone-related morbidity have examined for differing effects depending on time-of-year. See Kagawa et al. [1975, 1976], for instance.

[11] Studies examining the effects of ozone on individuals with chronic respiratory disease are summarized in Environmental Protection Agency [1978], esp. pp. 208-212.

[12] In clinical investigations of ozone health effects, lagged effects are of interest in part because they bear on the question of possible adaptation to ozone. See, for instance, Hackney _et al._ [1977]. See also Seskin [1979], pp. 281-283, and Seskin and Kawin [1981].

[13] See Theil [1971], pp. 607-615.

[13a] See U.S. Environmental Protection Agency [1978], especially pp. 249-250.

[14] A more sophisticated approach to ascertaining the effects of medical care on health may be found in Crocker _et al._ [1979] and Gerking and Schulze [1981]. See also Chappie and Lave [1982].

[15] See Goldsmith and Landaw [1968].

[16] For a recent example, see d'Arge, Kask, Case, Ben-David, Eubanks and Anderson [1982].

[17] They included Gilbert Omenn, Jane Koenig, Michael Morgan, Robert Frank, Alice Whittemore, and Michael Lebowitz. John Spengler also provided valuable advice.

[18] Schmidt [1976], pp. 40, 48.

[19] See Belsley, Kuh, and Welsch [1980].

[20] The following discussion draws heavily from the pioneering work of Belsley, Kuh, and Welsch [1980].

[21] The work of Mueller [1982] influences the following discussion.

References to Chapter 4

Bates, David, S. Field, M. Hazucha, C. Parent, and F. Silverman, "Pulmonary Function in Man After Short-term Exposure to Ozone," Archives of Environmental Health, vol. 27 (1973), pp. 183-188.

Belsley, D.A., E. Kuh, and R. Welsch, Regression Diagnostics (New York: Academic Press), 1982.

Chappie, Mike and Lester Lave, "The Health Effects of Air Pollution: A Re-analysis," Journal of Urban Economics, vol. 12 (1982), pp. 346-376.

Crocker, Thomas, William Schulze, and Shaul Ben-David, "Methods of Assessing Air Pollution Control Benefits: Vol. 1-Experiments in the Economics of Air Pollution Epidemiology," Environmental Protection Agency report no. EPA-600/5-79-001a, February 1979.

d'Arge, Ralph, Susak Kask, James Case, Shaul Ben-David, Larry Eubanks, and Curtis Anderson, "Air Pollution and Disease: An Evaluation of the NAS Twins," a report to the Environmental Protection Agency, Office of Research and Development, under contract no. CR-808893-01 to the Resource and Environmental Economics Laboratory, University of Wyoming, 1982.

Friedman, Robert, Sensitive Populations and Environmental Standards (Washington, D.C.: The Conservation Foundation), 1981.

Gerking, Shelby and William Schulze, "What Do We Know About the Benefits of Reduced Mortality from Pollution Control?" American Economic Review, vol. 71 (1981), pp. 228-234.

Goldsmith, John and Stephen Landaw, "Carbon Monoxide and Human Health," Science, vol. 162 (1968), pp. 1352-1359.

Graves, Philip and Ronald Krumm, Health and Air Quality (Washington, D.C.: American Enterprise Institute), 1981.

Hackney, J., C. Collier, W. Linn, and S. Mohler, "Adaptation to Short-term Respiratory Effects of Ozone in Men Exposed Repeatedly," Journal of Applied Physiology, vol. 43 (1977), pp. 82-85.

Kagawa, J. and T. Toyama, "Photochemical Air Pollution: Its Effects on Respiratory Function of Elementary School Children," Archives of Environmental Health, vol. 30 (1975), pp. 117-122.

______, ______ and M. Nakaza, "Pulmonary Function Tests in Children Exposed to Air Pollution," in A.J. Finkel and W.C. Duel (eds.), Clinical Implications of Air Pollution Research (Acton, Mass.: Publishing Sciences Group), 1976, pp. 305-320.

Krumm, Ronald and Philip Graves, "Morbidity and Pollution: Model Specification Analysis for Time-series Data on Hospital Admissions," Journal of Environmental Economics and Management, vol. 9 (1983), pp. 311-327.

McDonnell, W.F., D.H. Horstman, M.J. Hazucha, E. Seal, Jr., E.D. Haak, and S. Salaam, "Pulmonary Effects of Ozone Exposure During Exercise," Journal of Applied Physiology, vol. 54 (1983) pp. 1345-52.

Mueller, Charles F., The Economics of Labor Migration (New York: Academic Press, 1982).

Ostro, Bart, "The Effects of Air Pollution on Work Loss and Morbidity," Journal of Environmental Economics and Management, vol. 10 (forthcoming 1984).

Schmidt, Peter, Econometrics (New York: Marcek Dellker), 1976.

Seskin, Eugene, "An Analysis of Some Short-Term Health Effects of Air Pollution in the Washington D.C. Metropolitan Area," Journal of Urban Economics, vol. 6 (1979), pp. 275-291.

______ and Naomi Kawin, "Air Pollution and Health in Los Angeles," mimeo, December 1981.

Theil, Henri, Principles of Econometrics (New York: Wiley), 1971.

Whittemore, Alice and Edward Korn, "Asthama and Air Pollution in the Los Angeles Area," American Journal of Public Health, vol. 70 (1980), pp. 687-696.

Zagraniski, R., B.P. Leadered, and J.A. Stolwuk, "Ambient Sulfates, Photochemical Oxidants, and Acute Health Effects: An Epidemiological Study," Environmental Research, vol. 19 (1979), pp. 306-320.

Chapter 5

ALTERNATIVE OZONE CONCENTRATIONS AND CHANGES IN MORBIDITY: ILLUSTRATIVE ESTIMATES

In the introduction to this report, we pointed out that the original purpose of our project was to estimate where possible the dollar value of the changes in health associated with alternative ozone standards. Although our focus is on epidemiological findings in this report, the statistical results from the previous chapter can be used to estimate changes in morbidity that might result from different ozone concentrations. And these in turn are essential to economic benefit estimation.

This chapter illustrates how that might be done. In the next section we make some illustrative estimates of the "simple" reductions (increases) in acute and chronic morbidity that might result from reduced (increased) ambient ozone concentrations. These are termed simple because they do not take into account the possible feedback into acute morbidity of any ozone-related chronic illness. Those effects are addressed separately in Section 5.2. In Section 5.3 we attempt to put these illustrative calculations in some perspective by comparing the ozone-related changes in morbidity predicted from our models with those predicted to result from changes in other independent variables including smoking status, age, and income. Finally, in Section 5.4 we suggest how these predicted changes in morbidity might begin to be valued in a more fully developed and quantitative benefit-cost analysis.

5.1 Ozone-related Changes in Morbidity

Two points should be made clear at the outset of this exercise. First, because of their number we cannot calculate health improvements using every one of the models estimated in Chapter 4. Rather, we are forced to choose from among them when making illustrative estimates of the changes in acute and chronic morbidity that might accompany changes in ambient ozone concentrations. We have deliberately chosen from among those models where ozone concentrations were positively and significantly associated with acute or chronic illness.

This does not necessarily maximize the size of the estimated change in health status accompanying a hypothetical change in ozone, since larger but not significant coefficients can often be found among the models we have estimated. But it should be recalled that although the associations used in this chapter are all significant at at least the 10 percent level, the estimated coefficients on ozone--while virtually always positive--were insignificant as often as not in the analysis presented in Chapter 4. In fact, this was the rule rather than the exception in our investigation of ozone and chronic respiratory and cardiovascular disease. To repeat, we have deliberately chosen to illustrate the application of our estimates using models where positive and significant coefficient estimates were obtained.

The second point is of even greater importance. It deals with the most critical, but also the most often overlooked, caveat associated with the interpretation of regression results: these results can demonstrate correlations (or associations) but cannot prove causation. Thus, when one observes that ozone or smoking or some other regressor is positively and significantly associated with morbidity, one can say that this is consistent with the hypothesis that the relationship is causal, but it does not "prove" causation. This is important to what follows because using regression results to estimate changes in morbidity resulting from changes in ambient ozone can only be done with confidence if the relationship is causal. While clinical and toxicological studies have provided much evidence to suggest why and how such relationships could be causal, readers should keep in mind that our results are only consistent with this hypothesis.

5.1.1 Acute Morbidity

One of the areas where positive and significant associations were identified between ozone and acute morbidity was minor restricted activity days among adults, due to all causes. It is fairly straightforward to use several of the models finding this association to estimate changes in health status resulting from changes in ambient ozone. Acute Model 1 from Table 4-1 can be expressed as:

$$(1) \quad \text{MRADs} = .772 + 2.48\ \text{03NR01} + \sum_i \hat{\beta}_i X_i \qquad \text{, where}$$

MRADs = number of minor restricted activity days due to all causes for a given individual during the two-week reference period

03NR01 = average daily maximum 1-hour ozone concentration during that period (in ppm) at the monitor nearest the individual's home

$\hat{\beta}_i X_i$ = the other independent variables and their estimated coefficients.

Differentiating (1) with respect to ozone, one gets:

$$(2) \quad \partial(\mathrm{MRAD}_s)/\partial(\mathrm{03NR01}) = 2.48$$

In other words, a one part-per-million increase (decrease) in the average daily maximum 1-hour ozone concentration during a two week reference period would mean 2.48 additional (fewer) minor restricted activity days due to all causes during that period for each individual. A more meaningful way to interpret this effect results from multiplying 2.48 by 10^{-2} since ozone concentrations are usually expressed in hundredths of parts-per-million. Thus each .01 ppm change in the average daily maximum would result in .0248 additional or fewer MRADS per person during a two-week period. Multiplying by 26 to convert this to an annual figure suggests that each .01 increase (decrease) in the average daily maximum 1-hour ozone concentration will result in 0.64 additional (fewer) MRADs due to all causes per person-year. Extrapolating to a population of about 110 million adults (persons $\geq$ 17 years) in SMSAs in the United States, this implies 70.4 million additional (fewer) all-cause minor restricted activity days per year if the average daily maximum ozone concentration was increased (decreased) by 0.01 ppm in each of these SMSAs.

This estimate can be compared with one derived using another of the models we estimated. Consider, for instance, Acute Model (5) from Table 4-1, where ozone was measured by the square of the average daily maximum (the coefficient on the ozone variable had a higher t-statistic in this model than in Model 1). Using Model 5, the partial derivitive is:

(3) $\partial(MRAD)/\partial(O3NR01) = 2 \times 16.7 \times O3NR01$

Evaluating ozone at the mean of the sample (.045), the change in the number of all-cause MRADs per person-year for each 0.01 ppm change in the average daily maximum reading is:

(4) $2 \times 16.7 \times 0.045 \times .01 \times 26 = 0.39.$

Applying this figure to the same adult SMSA population used above yields an estimated change in all-cause MRADs of 42.9 million per year for a 0.01 ppm change in the average daily maximum 1-hour ozone concentration.

The difference between these two illustrative estimates is due simply to the difference in the two estimated coefficients. An examination of the ozone coefficients in Table 4-1, where these two models of acute mordibity were originally presented, indicates that still different estimates would have been obtained had alternative models been selected for these calculations.

Another area where we often found ozone to be positively and significantly associated with morbidity was acute respiratory disease among adults. We can make estimates similar to those above using models in which the dependent variable is the number of total restricted activity days due to respiratory disease during the two week reference period. (Recall we found no statistically significant association between ozone and respiratory-related work loss or bed disability among adults.) An example is Model 48 where the coefficient on ozone is positive and significant at the 5 percent level. From that equation, it can be seen that:

(5) $\partial(\text{TRADRSP})/\partial(\text{O3NR01}) = 1.54$

Thus, to obtain an estimate of the number of additional total restricted activity days per person-year due to respiratory disease for a 0.01 ppm change in the average daily maximum ozone concentration, one makes the following simple calculation: $1.54 \times .01 \times 26 = 0.39$. A .01 ppm change in the average daily maximum 1-hour concentration of ozone would therefore imply a change of 42.9 million total restricted activity days due to respiratory disease alone each year, assuming an adult SMSA population of 110 million in the United States.

5.1.2 Chronic Morbidity

The results in Chapter 4 can also be used to estimate the effect of a change in the average annual hourly ozone concentration (as opposed to the average daily maximum) on the likelihood of chronic respiratory disease (CRD). In this exercise, it is useful to work with the elasticity of the probability of CRD with respect to ozone concentrations, defined as the percentage change in the former for a given percentage change in the latter. Consider first the general form of the estimating equation used in our analysis of CRD.

It is:

(6) $$\ell n\left(\frac{P_{CRD}}{1-P_{CRD}}\right) = \alpha + \sum_{i=1}^{n} \beta_i X_i$$ where

P_{CRD} = the probability that an individual has a CRD

$X_1, \ldots, X_n$ = ozone and the other independent variables.

Defining $\eta_{P,03}$ as $(\partial P_{CRD}/\partial(03))\ (03/P_{CRD})$, we can differentiate (6) to get:

(7) $$\eta_{P,03} = (1-P_{CRD})\hat{\beta}_1\ (\overline{03})$$ where

$\hat{\beta}_1$ = the estimated coefficient on ozone in the regression explaining P_{CRD}.

$\overline{P}_{CRD}$, $\overline{03}$ = the sample means of P_{CRD} and 03.

Using chronic Model 9 from Table 4-14 to illustrate, it can be seen that when (7) is evaluated at the means of ozone and the probability of CRD, then $\eta_{P,03} = (.92)(13.9)(.02) = 0.25$, since $\overline{P}_{CRD} = .08$ and $\overline{03} = .02$ in the sample over which Model 9 was estimated. Using this to estimate changes in the numbers of individuals with CRD, consider a 10 percent increase (decrease) in the average annual hourly ozone concentrations in the SMSAs of the United States. Since $\eta_{P,03} = 0.25$, this implies a 2.5 percent increase (decrease) in the probability of CRD, or an increase (decrease) from approximately 8 percent to 8.2 (from 8 to 7.8 percent) percent of the population. Since there are roughly 110 million adults in SMSAs in the United States, this implies an eventual increase (decrease) of 220,000 cases of chronic respiratory disease at a given point in time. Depending upon the model used to determine the coefficient on ozone, this estimate could be higher or lower. For instance, if chronic Model 16 from

Table 4-14 had been used, the estimated change in numbers of cases of CRD for any change in ozone would be higher by about 60 percent. On the other hand, using coefficients from models 10 or 13 would cut the estimate above in half.

To summarize, if the coefficients used in the analysis in this and the previous section accurately reflect the association between ozone and acute and chronic respiratory and other disease, and if the relationships are causal, they imply potentially significant changes in acute and chronic morbidity for changes in ambient ozone concentrations. For instance, a hypothetical change of 0.01 ppm in average daily maximum 1-hour ozone concentrations implies a change of several tens of millions in the number of minor restricted activity days due to all causes among adults, and a similar change in the number of total restricted activity days due to acute respiratory disease. A 10 percent change in the longer-term ozone concentration, measured by average annual hourly readings, implies an eventual 2.5 percent reduction in the likelihood of CRD, or about 220,000 cases at 1980 population levels.

5.2 Acute and Chronic Illness: Feedback Effects

Even if ambient ozone had no direct effect on acute morbidity, it might influence it indirectly via its possible effect on chronic illness, which our results suggest is strongly associated with acute disease. For instance, in acute Model 1 from Table 4-1 (used above to estimate changes in all-cause MRADs due to changed ozone levels) the coefficient on the dummy variable indicating an individual has a chronic illness is positive and significant at greater than the 1 percent level. It is useful to take these possible effects into account in making illustrative estimates.

A simple model may illustrate this point. Suppose

(8) $S = S^A + S^C$,

where S, S^A, and S^C represent respectively total days spent sick during some period, days spent sick due to acute illness, and days ill due to a chronic illness (recognizing that the distinction between acute and chronic illness is not a simple one to make). Suppose further that:

(9a) $S^A = f(X,Z,S^C)$ and

(9b) $S^C = g(Y,Z)$, where

X and Y are personal, socioeconomic, or non-ozone environmental characteristics that predispose individuals to acute and chronic illness respectively, and Z is a measure of ambient ozone which is assumed to affect both acute and chronic illness. (In actuality, acute illness is probably related more to peak than average concentrations, while chronic disease may be more influenced by cumulative dose and therefore more closely related to average annual concentrations. We could link peak and average concentrations through some complex relationship in this model but choose the present formulations for simplicity.)

Differentiating (8) with respect to Z, we obtain:

(10) $\partial S/\partial Z = f_Z + g_z(f_{S^C} + 1)$.

The term f_Z in (10) is the direct effect of ozone on acute morbidity, as estimated in our models of MRADs, TRADs, WLDs, and BDDs in Chapter 4. The second term, $g_Z(f_{S^C} + 1)$, represents the potential direct and indirect effects of ozone on total sickness operating through chronic illness.

For an illustrative quantitative estimate of the magnitude of the combined effect, note that in acute Model 1 a 10 percent increase in the average daily maximum 1-hour ozone concentration (from a mean value of .045) would add to each individual's all-cause MRADs during a two-week period in the following way:

(11) $\partial(\mathrm{MRAD})/\partial(\mathrm{03NR01}) = .0248 \times .0045 = .011$

Thus, each adult SMSA resident would suffer a direct increase in MRADs of 0.29 per year (=26 x .011). For an adult SMSA population of 110 million, this would amount to an additional 31.9 million all-cause MRADs per year.

If average annual hourly ozone concentrations also increased by 10 percent, then from chronic Model 9 in Table 4-14 it can be determined that the likelihood of chronic respiratory disease will increase by 2.5 percent, since $\eta_{P,03} = 0.25$. As demonstrated above, this implies an eventual increase of 220,000 cases of chronic respiratory disease each year. From acute Model 1 (Table 4-1), the effect of this change in chronic illness on acute morbidity as measured by all-cause MRADs can be determined. The coefficient on CHRLMDUM (the dummy variable indicating the presence of a chronic illness) is approximately 1.30, indicating that a chronic illness induces 33.8 additional all-cause MRADs each year (26 x 1.3). Since each of 220,000 individuals would be thus affected, MRADs would increase by about 7.4 million each year among the adult SMSA population in the United States. In other words, the indirect effect on all-cause MRADs of an ozone-induced increase in the likelihood of chronic respiratory disease would amount to about 23 percent ($7.4 \times 10^6 \div 31.9 \times 10^6$) of the direct effect. This points out the importance of considering the possible effects of ozone on chronic morbidity.

Again we must remind the reader that the positive and significant association identified in Model 9 (table 4-14) and used for these illustrative estimates is not a robust one. The coefficients on ozone in most other similar models of the determinants of CRD are not significant (although always positive). We use Model 9 here only to illustrate the possible magnitude of a "feedback" effect if one exists, and we abstract from the fact that chronic illness would take some time to develop or abate if ozone concentrations were to change. Our results suggest that the effect may be a significant one if further research confirms the likelihood of its existence.

5.3 Relative Risks

There are a number of ways one can assess the credibility of empirical results. First, t-ratios or other measures of statistical significance shed light on the likelihood that observed associations are accidental. Next, within a particular regression, it is important that other variables behave plausibly. To the extent that many other variables have counterintuitive signs, it casts doubt on the validity of those estimated effects that do conform to prior expectations. Also, the sensitivity of observed associations must be assessed with respect to the measurement of dependent and independent variables, the inclusion or exclusion of certain regressors, the functional form of the estimating equation, the presence of "influential" observations, and other changes in the basic approach to estimation. In Chapter 4 we relied on all these means to determine the plausibility of our findings.

There is another approach to help determine whether observed associations seem plausble, one which might be referred to as relative risk analysis. Under this approach, the predicted effects on health of changes in ozone (or other variables of interest) are compared with the predicted effects of changes in other important regressors. If considerable information is available about these other risk factors, such a comparison can sometimes suggest the plausibility of the observed associations between the measure of health and the variable(s) of interest. In this section, we briefly explore several of the models estimated in Chapter 4 and used in this chapter for illustrative health benefit calculations.

Consider first Model 1 from Table 4-1 above. Using this model, it was estimated that an increase (decrease) of 0.01 ppm in the average daily maximum 1-hour ozone concentration during the two week reference period would result in 0.0248 additional (fewer) all-cause MRADs during the two-week period (or 0.64 more or less MRADs per person-year). How does this compare with the effects of changes in other independent variables? According to Model 1, a change in smoking status--from being a never- or former smoker to being a current or occasional smoker--would add 0.15 to the number of all-cause MRADs during a two week period (a finding significant at the 5 percent level). This would be equivalent to 3.9 additional all-cause MRADs per person-year, or six times the effect of a 0.01 ppm increase in the average daily maximum ozone reading.

It is not surprising that the hypothetical change in ozone is predicted to have a smaller effect on acute illness than a change in smoking status, although one might have expected the relative size of the ozone effect to be somewhat smaller. That it is not may be due to the crudeness of the smoking dummy variable, discussed at some length in Chapter 4. By contrast to the effect of smoking, Model 1 from Table 4-1 suggests that the presence of a chronic limitation would add 1.3 all-cause MRADs to each two-week period, or an additional 33.8 person-year (significant at the 1 percent level). This effect is more than 50 times the effect of a .01 ppm increase in ozone. If Model 5 (Table 4-1) is used to make such comparisons, the difference between a change in smoking status and a 0.01 change in the average daily maximum ozone concentration is larger. In this case, a change in smoking status again adds 3.9 all-cause MRADs to an individual's yearly total, as compared to an additional 0.39 from a hypothetical 0.01 ppm change in ozone. Thus, the predicted difference is an order of magnitude.

Turning to acute respiratory disease, Model 48 (Table 4-7) can be used as a basis for comparing relative risks as estimated in this study. According to that model, a 0.01 ppm increase in average daily maximum ozone concentrations would result in an additional 0.0154 total restricted activity days per person due to respiratory disease during a two-week period, or about 0.4 per person-year. Model 48 also suggests that a change in smoking status like that discussed above (in which SMOKY1N0 takes on the value 1 rather than 0) would add 0.03 TRADs per person due to respiratory disease to each two-week period, or 0.78 per person-year. In this example, the change in smoking status is only twice as harmful as a 0.01 ppm change in ozone, even though the explicit focus in Model 48 is respiratory disease--as opposed to acute morbidity due to all causes as in Models 1 and 5. Using Model 48, a chronic illness is predicted to have sixteen times the effect on respiratory-related TRADs as the illustrative 0.01 ppm change in ozone.

It is useful to perform this same experiment to put in some perspective our findings about ozone and chronic respiratory disease. To do so, consider Model 9 from Table 4-14 as a basis for such a comparison, a model which relates each of the independent variables to the probability that an individual will have a chronic respiratory disease. In this as well as in the other models explaining chronic illness, ozone is measured by the average annual hourly concentration rather than by the average of the highest daily 1-hour readings, which were used in the analysis of acute morbidity. This is important because for the 14,441 adults that constitute our overall sample, the average annual hourly reading in 1979 was 0.02 ppm. Thus, when we discuss a hypothetical 0.01 ppm change in the average annual hourly reading, it amounts to a 50 percent change from the mean. Since the average daily maximum value for these same 14,441 people was 0.045 during their reference period, the 0.01 ppm change referred to previously in this section was only slightly more than 20 percent of the mean.

According to Model 9, an increase of 0.01 ppm in the average annual hourly ozone concentration would increase the probability of CRD by 0.0103 percentage points $[\partial P/\partial(\mathrm{03ANARAV}) = \overline{P}_{CRD}(1-\overline{P}_{CRD})\ \hat{\beta} = (.08)(.92)(13.94)(.01) = .0103]$. A change of one cigarette smoked per day would increase the probability of CRD by .00138 [$\partial P/\partial(\text{NCIGSDYN}) = (0.8)(.92)(.0187) = .00138$]. Thus, an individual would have to smoke an additional 7.5 cigarettes per day to incur the same added risk of CRD as is predicted to arise from an increase of 0.01 ppm (or 50 percent) in the average annual hourly ozone concentration. However, it should be observed that the coefficient on NCIGSDYN used to make this illustrative comparison is not significant at conventionally accepted levels (its t-ratio is only 0.51).

Using Model 9, it can also be seen that approximately twenty years would have to be added to the age of an average individual to have the same effect on the probability of CRD as would be associated with a 0.01 ppm change in ozone concentrations. It can also be determined from Model 9 that an increase of $5720 in an individual's annual household income would have the same effect on the likelihood of CRD as a decrease of 0.01 ppm ozone, measured by average hourly concentration.

Model 16 in Table 4-14 is another in which ozone was positively and statistically significantly associated with the probability of CRD. Using it as the basis for a comparison of relative risks, an extremely implausible result obtains. From that model, an increase of 0.01 ppm in the average annual hourly ozone concentration is predicted to increase P_{CRD} by 1.66 percentage points. An equivalent increase due to smoking would require an individual to smoke about 59 additional cigarettes per day, an increase of three packs (recall this is based on a parameter estimate that was not significant). In spite of its representing a 50 percent increase in the average annual hourly ozone concentration, it is obviously implausible that an increase of 0.01 ppm would increase the likelihood of chronic respiratory disease by the same amount as an increase of three packs of cigarettes smoked each day. Model 16 also suggests that the hypothetical 0.01 ppm increase considered here would be equivalent, in terms of increasing the likelihood of CRD, to adding forty years to the age of the average individual or subtracting $11,000 from annual household income (the coefficients on both age and income were significant at the 5 percent level in Model 16). Table 5-1 presents these relative risks.

To summarize, in the case of acute morbidity due to all causes (as measured by MRADs), our calculations suggest that approximately a 20 percent increase in the average daily maximum ozone concentration (from a mean of about 0.045 ppm) will add about one-sixth to one-tenth as many MRADs to an individual's predicted total as a change in that individual's smoking status (from never or former smoker to current or occasional smoker). The coefficients on both ozone and smoking status were significant and had the expected signs in the Models on which these estimates were based. While we might have expected the relative risk associated with smoking to be larger, the smoking variable we employed left much to be desired. It may bias downward the predicted effect of smoking on health, and thus affect our comparisons of relative risks.

Shifting to chronic respiratory disease and its correlates, we found using one model that an increase in the average annual hourly ozone concentration of 0.01 ppm (a 50 percent change from the mean) was equivalent to smoking an additional 7 1/2 cigarettes per day in terms of

Table 5-1. Changes in Selected Independent Variables Required to Change the Probability of Chronic Respiratory Disease by the Same Amount as a Change of 0.01 ppm in the Average Annual Ozone Concentration

Variable (and description)	Change Required[1]
o NCIGSDYN (Number of cigarettes smoked per day)	59
o AGE (age in years)*	40
o INCOMCON (annual household income in dollars)*	11,000

---

*Coefficient significant at 95 percent level.

[1] Using Model 16 (Table 4-14).

increasing the likelihood of CRD. Using another model in which ozone was positively and significantly associated with CRD, we found that an individual would have to smoke an additional three packs per day to increase his or her risk of CRD by the same amount as it would be affected by an increase of 0.01 ppm. This latter result is highly implausible. It serves as a reminder of the problems inherent in selecting somewhat unrepresentative results to make these sorts of comparisons, not to mention making extrapolations based on insignificant parameter estimates.

5.4 Monetary Benefit Estimates

Although we do not derive monetary estimates of health benefits in this report, the epidemiological findings in Chapter 4 and the illustrative calculations in the previous sections of this Chapter could be used for that purpose. In this section we suggest how. This section is no more than suggestive, however. RFF is just beginning work on the identification and valuation of the benefits and costs of alternative ozone standards. More rigorous analysis of human health benefits, possibly using the present epidemiological study as well as many other epidemiological, clinical, and/or toxicological studies, will go into the valuation issue in much greater detail.

The NAAQS for ozone is expresed in terms of the maximum allowable concentration (0.12 ppm) for the second highest hourly reading at a particular monitor in a given year (actually the standard is not to be exceeded more than three times in three years on an expected value basis). However, in our analysis of acute and chronic morbidity, ozone was measured, respectively, by the average of the daily 1-hour maxima during the two-week reference period, and by the annual hourly concentration averaged over 1979 or the period 1974-79. Thus, the first step in estimating monetary health benefits using the dose-response relationships in this report would involve linking ozone as currently measured for the standard to the way it is characterized in our study.

There are a number of ways this could be done, only two of which are discussed here. Consider first a simple "proportional rollback" approach. Under such an approach, if a hypothetical change in the ozone standard from 0.12 ppm to 0.14 ppm (or about 16.7 percent) were under consideration, one would assume that all ozone readings would be shifted up by the same percentage. Thus, if the average daily maximum 1-hour reading was currently 0.06 ppm in a particular area, the new value would be assumed to be 0.07 ppm. This increment of 0.01 ppm could then be used to predict additional all-cause MRADs or respiratory TRADs in the case of acute illness. Similarly, if the average annual hourly concentration in an area was 0.024 ppm prior to the hypothetical change, it would be assumed to be 0.028 after the relaxation. This change would then be used to estimate the altered probability of chronic respiratory disease if it was determined to use our ambiguous findings about ozone and CRD as the basis for a benefit calculation. The same exercise could be repeated for a hypothetical tightening of the ozone standard, say from 0.12 ppm to 0.10 ppm. Again, a uniform percentage change could be assumed.

A more complicated and sophisticated approach could also be taken. This would involve statistical estimation of the relationship between the second highest hourly reading at a given monitor during a given year and the two measures of ozone used in our analysis--the average daily maximum (in the case of acute morbidity) and the average annual hourly concentration (in the case of chronic morbidity). Using the 1979 hourly data from all the ozone monitors meeting certain reporting criteria, several higher-order polynomials as well as other functional forms could be estimated to determine which best fits the data. The resulting equation could then be used to predict changes in the ozone concentrations used for the dose-response estimation given changes in the standard.

This would have one advantage compared to the proportional rollback approach: it would allow for the likely possibility that increases or decreases in allowable extreme values (like the second highest hourly reading in a year) would not imply identical percentage changes throughout the whole distribution of hourly values. If extreme values are to be reduced, for example, it is possible that the mean of the distribution--the average annual hourly average in this case--would be changed very little. This more sophisticated conversion technique would be a more realistic means of anticipating the efects of a change in the national ambient air quality standard.

Other than to point out that predicted changes in air quality must then be linked to exposed populations in order to determine health effects, we say nothing here about population exposures. Determining populations at risk depends critically on the assumptions made in a benefits analysis--for example, if standards are tightened, do areas currently experiencing air quality better than the new standards "drift up" to the new standards? Or, do more stringent standards only affect the ultimate air quality of areas currently just meeting or exceeding the old standards? Questions like these will be addresed in a subsequent analysis. Suffice it to say they must be addressed if dose-response relationships like those above are to be used to make monetary estimates of ozone-related health benefits.

Once dose-response relationships are coupled with populations at risk, the final step in estimating benefits in monetary terms is assigning dollar values to the various types of resulting health changes. Here again there are a number of ways this could be accomplished. All are linked in one way or another to an individual's willingness to pay (or be compensated) for an improvement (or worsening) in his or her health status. This is because willingness to pay is the conceptually correct measure of economic benefit in welfare economics.[1]

Direct elicitation of individuals' willingness to pay for short-term improvements in health has been attempted through questionnaires.[2] Individuals were asked how much they would pay for the alleviation of symptoms such as shortness of breath, chest pain, coughing and sneezing, and head congestion and closely related symptoms. The duration of the reduction in symptoms was allowed to vary in the study so that estimates were obtained for one day's worth of alleviation, as well as one week's and finally three months' worth of relief. (Thus, the resulting valuations would be most useful in preparing monetary estimates of changes in acute as opposed to chronic illness.) Willingness to pay was elicited for both "mild" and "severe" cases in each symptomatic category. Average willingness to pay of those surveyed varied between $2.31 to avoid one day of minor coughing and sneezing to $97.80 to avoid three months of severe shortness of breath (at what appear to be 1975 prices and incomes).

Like all estimates obtained in surveys, these are subject to hard questions about sample selection bias, as well as bias resulting from the way the survey was designed and administered. We mention these estimates here because the symptoms for which valuations were elicited are similar to some of those thought to result from exposure to peak concentrations of ozone. Therefore, they might be useful in developing a range of values for the health changes obtained using the dose-response relationships estimated in the "acute" portion of our study.

Concerning our estimates of the effect of ozone on acute morbidity, an interesting valuation problem arises. Recall that we found no relationship between ozone and either bed disability days or work or school loss days among the adults and children analyzed. But we did find frequent positive associations between ozone and minor restricted activity days due to all causes and total restricted activity due to respiratory disease (in adults). Yet these restricted activity days may be more difficult to value than the more serious types of acute morbidity because they are more difficult to conceptualize and define.

For instance, an individual might be willing to pay very little to avoid an acute illness that prevents him from doing a minor but necessary and uninteresting chore, yet would be willing to pay much more if this same illness or condition prevented him from taking part in an eagerly anticipated social or recreational event. Yet both would be considered restrictions in regular activity under the HIS protocol. This suggests that even survey data on willingness to pay would be difficult to interpret and use unless questions were tied to specific types of restricted activity. Moreover, since restrictions on activity as defined by HIS will generally not entail out-of-pocket spending or lost earnings, even this cost-of-illness proxy for willingness to pay will be of little assistance. These factors suggest that a range of values should be used in converting restricted activity days to monetary benefits.

Difficulties arise in valuing chronic illness, as well, although they are not of as much consequence for this report because of the very tentative nature of our findings regarding ozone and chronic respiratory disease. There are some survey data on willingness to pay to avoid chronic illness comparable to those cited above for acute morbidity.[3] These data, which suggest that individuals would be willing to spend as much as several hundred dollars per year to be free of chronic illness related to air pollution, must be treated very carefully--individuals can be expected to have much more difficulty assigning meaningful values to long periods of time free from chronic respiratory or other kinds of disease than to a day or week free from coughing and sneezing. Among other things, such estimates would depend heavily on the prior health histories of individuals surveyed. In addition, estimates will be sensitive to the sample upon which they are based, as well as upon a number of the characteristics of the survey itself and the way it is administered.

Estimates have been made of the annual costs associated with various types of chronic illness. One study reported the annual costs associated with CRD comprised of (i) a "direct" component consisting of doctors' charges plus the cost of hospitalization and/or nursing home care, (ii) an "indirect" component reflecting lost earnings, and (iii) a component called "indirect mortality cost" which is the product of the probability of death from the disease times the earnings that would be lost in the event of death.[4] By disease, these annual costs were estimated to be: chronic bronchitis, $237; bronchiectasis, $701; emphysema, $1294; chronic intestitial pneumonia, $195 (all in $ 1981).

For a number of reasons these estimates would not be particularly good proxies for individuals' willingness to pay to reduce the prevalence of the diseases considered in our study. First, like all cost-of-illness estimates, they ignore both the pain and suffering associated with illness as well as any defensive measures individuals may have taken to protect themselves against chronic illness.[5] Second, to the extent that "indirect mortality cost" is an appropriate component of such an estimate--and this is unclear--it should be based not on foregone earnings during the extended years of life, but rather on willingness to pay for reduced mortality risks. Nevertheless, if used with care these and other recent estimates of the costs associated with chronic respiratory disease might be of some use in making monetary estimates of the health benefits associated with alternative NAAQSs for ozone.

Footnotes to Chapter 5

[1] See Just, Hueth, and Schmitz [1982] or Freeman [1979].

[2] For example, see Loehman *et al.* [1979].

[3] See Brookshire, d'Arge, Schulze, and Thayer [1979] and Loehman, Boldt, and Chaikin [1981].

[4] See d'Arge, Kask, Case, Ben-David, Eubanks, and Anderson [1982], pp. 39-43.

[5] See Appendix C to this report for an analysis of the effects such omissions can have on benefit estimates.

References to Chapter 5

Brookshire, David, Ralph d'Arge, William Schulze, and Mark Thayer, "Methods Development for Assessing Pollution Control Benefits," Volume II, "Experiments in Valuing Non-market Goods: A Case Study of Alternative Benefit Measures of Air Pollution Control in the South Coast Air Basin of Southern California," Publication no. EPA-600/5-79-001b, February 1979.

d'Arge, Ralph, Susak Kask, James Case, Shaul Ben-David, Larry Eubanks, and Curtis Anderson, "Air Pollution and Disease: An Evaluation of the NAS Twins," a report to the Environmental Protection Agency, Office of Research and Development under contract no. CR-808893-01 to the Resource and Environmental Economics Laboratory, University of Wyoming, 1982.

Freeman, A. Myrick, The Benefits of Environmental Improvement (Baltimore, Md.: Johns Hopkins University Press for Resources for the Future), 1979.

Just, Richard, Darrell Hueth and Andrew Schmitz, Applied Welfare Economics and Public Policy (Englewood Cliffs, N.J.: Prentice-Hall), 1982.

Loehman, E., S.V. Berg, A.A. Arroyo, R.A. Hedinger, J.M. Schwartz, M.E. Shaw, R.W. Fahien, V.H. De, R.P. Fishe, D.E. Rio, W.F. Rossley, and A.E. Green, "Distributional Analysis of Regional Beneits and Costs of Air Quality Control," Journal of Environmental Economics and Management, vol. 6 (1979), pp. 222-243.

Loehman, Edna, David Boldt, and Kathleen Chaikin, "Measuring Benefits of Air Quality Improvements in the San Francisco Bay Area," report prepared by SRI International for the Department of Economics, University of Wyoming, May 1981.